CREOLE LANGUAGE LIBRARY VOLUME 29

STRUCTURE AND VARIATION IN LANGUAGE CONTACT

EDITED BY
Ana Deumert
Stephanie Durrleman

JOHN BENJAMINS PUBLISHING COMPANY

Structure and Variation in Language Contact

Creole Language Library (CLL)

A companion series to the "*Journal of Pidgin & Creole Languages*"

Editor

John Victor Singler
New York

Editorial Advisory Board

Mervyn Alleyne
Kingston, Jamaica

Marlyse Baptista
Athens, Georgia

Lawrence Carrington
Trinidad

Glenn Gilbert
Carbondale, Illinois

George Huttar
Dallas

John Holm
Coimbra

Silvia Kouwenberg
Kingston, Jamaica

Salikoko Mufwene
Chicago

Pieter Muysken
Nijmegen

Peter Mühlhäusler
Adelaide

Norval Smith
Amsterdam

Sarah G. Thomason
Ann Arbor, Michigan

Volume 29

Structure and Variation in Language Contact
Edited by Ana Deumert and Stephanie Durrleman

Structure and Variation in Language Contact

Edited by

Ana Deumert

University of Cape Town, Monash University

Stephanie Durrleman

University of Geneva

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

Structure and variation in language contact / edited by Ana Deumert, Stephanie Durrleman.
p. cm. (Creole Language Library, ISSN 0920–9026 ; v. 29)
Includes bibliographical references and index.
1. Languages in contact. 2. Language and languages--Variation.
I. Deumert, Ana. II. Durrleman, Stephanie.

P130.5.S75 2006
417'.7--dc22 2006049872
ISBN 90 272 5251 3 (Hb; alk. paper)

© 2006 – 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 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

In memory of Jacques Arends

Table of contents

Introduction 1
Ana Deumert and Stephanie Durrleman

Part I. Structure

The phonetics of tone in Saramaccan 9
Jeff Good

Tracing the origin of modality in the creoles of Suriname 29
Bettina Migge

Modeling Creole Genesis: Headedness in morphology 61
Tonjes Veenstra

The restructuring of tense/aspect systems in Creole formation 85
Donald Winford

Syntactic properties of negation in Chinook Jargon, with a comparison to two source languages 111
Zvjezdana Vrzić

Sri Lankan Malay morphosyntax: Lankan or Malay? 135
Peter Slomanson

Sri Lanka Malay: Creole or convert? 159
Ian Smith and Scott Paauw

The advantages of a blockage-based etymological dictionary for proven or putative relexified languages (Extrapolating from the Yiddish experience) 183
Paul Wexler

Part II. Variation

A fresh look at habitual *be* in AAVE 203
Chris Collins

Oral narrative and tense in urban Bahamian Creole English 225
Stephanie Hackert

Aspects of variation in educated Nigerian Pidgin: Verbal structures 243
Dagmar Deuber

A linguistic time-capsule: Plural /s/ reduction in Afro-Portuguese and Afro-Hispanic historical texts 263
Fernanda L. Ferreira

The progressive in the spoken Papiamentu of Aruba 291
Tara Sanchez

Was Haitian ever more like French? 315
Mikael Parkvall

The late transfer of serial verb constructions as stylistic variants in Saramaccan creole 337
Marvin Kramer

Index 373

Introduction

Ana Deumert and Stephanie Durrleman
University of Cape Town, Monash University (South Africa, Australia) / University of Geneva, Switzerland

The fifteen papers in this collection were presented at the SPCL meetings in Atlanta, Boston and Hawai'i in 2003 and 2004, and reflect – from various perspectives and using different types of data – on the interplay between structure and variation in contact languages, both synchronically and diachronically. The authors show that language contact is a complex process of language change, involving both linguistic innovation (as a result of L2A and L1A), and the reanalysis of super- and substrate input within historically specific contact ecologies.

A need to question and test existing hypotheses regarding pidginization/creolization is evident in all contributions. Among the claims discussed in these papers are the following:

(a) the idea that adults innovate while children regulate (Veenstra);
(b) the assumption that most structures in contact languages can be explained as continuities from either superstrate or substrate language(s) (Veenstra, Winford, Good, Migge, Vrzić);
(c) the hypothesis that the creole prototype (as proposed, e.g., by Bickerton 1981 and McWhorter 1998) is a form of language which does not show marked or 'ornamental' structures (Good, Kramer). This, as argued by Smith and Paauw, might well be a model which is "biased towards Atlantic creoles with [highly heterogeneous] West-African substrates".
(d) Lefebvre's (1998) relexification hypothesis (Migge);
(e) the importance of decreolization in explaining variation and language change in contact languages (Collins, Deuber, Parkvall).

The papers call attention to the need for careful sociolinguistic data collection, including various discourse types and social groups. Authors draw on a variety of sources, e.g., informal English as represented on the internet (Collins), historical texts (Ferreira, Kramer), sociolinguistic interviews (Hackert, Sanchez, Deuber), and elicitation (Good, Slomanson). The papers provide ample evidence for the observation that pidgin/creole studies is today a mature subfield of linguistics, and is making important

contributions to general linguistic theory. Good, for example, argues in his paper on tone in Saramaccan that the typologically unique synchronic situation of a lexical split between a class of words marked for accent and a class marked for tone (i.e. Saramaccan shows a split tonal/accentual phonology) might be the result of a logically possible (but otherwise unattested) contact phenomenon between European accent languages and African tonal languages. Rather than 'levelling' its lexicon towards an African type or a European type, Saramaccan appears to have maintained two parallel prosodic systems, one, roughly speaking, with 'European' characteristics and another with 'African' characteristics. That language contact can lead to linguistic structures which are typologically unique (and often unpredictable) is also evident in Slomanson's paper on Sri Lankan Malay, a language which, according to his analysis, seems "a compromise between contact Malay and (Sri Lankan Muslim) Tamil", and structural parallels between Tamil and Sri Lankan Malay remain incomplete and "asymmetrical across syntactic domains".

The contributions to this volume consider a range of languages. Several of the papers analyze data from Surinamese creoles (Veenstra, Winford, Good, Kramer, Migge, Winford). Other contact varieties discussed include: Chinook Jargon (Vrzić), Yiddish (Wexler), AAVE (Collins), Haitian Creole (Winford, Parkvall), Afro-Hispanic and Afro-Portuguese varieties (Ferreira), ('educated') Nigerian Pidgin (Deuber), Sri Lankan Malay (Smith & Paauw, Slomanson), Papiamentu (Sanchez), and Bahamian Creole English (Hackert). The authors also provide analyses for a wide variety of grammatical structures: VO-ordering and affixation (Veenstra), agglutination (Parkvall), negation (Vrzić, Deuber, Slomanson), TMAs (Collins, Winford, Deuber, Slomanson, Sanchez, Hackert, Smith and Paauw, Migge), plural marking (Ferreira), the copula (Deuber), and serial verb constructions (Kramer).

The importance of substrate structures is noted in many of the articles. For example, sentential negation in Chinook Jargon is shown to reflect syntactic features found in the two most relevant source languages, Lower Chinook (Chinkookan) and Upper Chehalis (Salishan). The properties of negation in Chinook Jargon can thus be explained as the result of substrate influence and transfer, a process that was supported by the 'structural congruence' of the dominant contact languages (Vrzić). Smith and Paauw discuss TMA systems. Although similarities between Sri Lankan Malay and Tamil with respect to TMA markers in Tamil and Sri Lankan Malay are structurally congruent, their lexical expression of these categories is different (with Sri Lankan Malay drawing on contact or vehicular Malay as a lexifier language). Yet, at the same time, TMA marking in Sri Lankan Malay is not simply a calque of Tamil since some main verbs allow for optional tense marking (which is not possible in Tamil), a likely 'carryover' from contact or vehicular Malay. Similarly, Migge shows that the TMA systems in the creoles of Suriname show input from both the Gbe language and the European superstrate: although "many aspects of the creole modality system have their source in Gbe language … it is clear that they are in no way exact (or in some cases even close) replicas of the Gbe modality systems" – they show not only lexical but also semantic and syntactic influence from the European superstrate. Finally, Veenstra, and

also Winford, draw attention to the fact that creolization is more than simply a process of reanalysis of substrate and/or superstrate structures, but involves innovations which have their basis in both L2A (adults) and L1A (children) strategies. Veenstra shows that, in the case of the Surinamese creoles, the structure of affixation is "radically different from those that are found in the languages that were present in the original contact situation". Affixation thus cannot be explained from the perspective of either superstrate or substrate influence. Instead, the structure of affixation in the Surinamese creoles is an innovation which took place during nativization (i.e. in this case children were the innovators).

Many of the papers pay detailed attention to the need for the careful analysis of patterns of linguistic variation to understand and explain the outcomes of contact-induced language change:

– Collins presents a fresh perspective on the diachronic evolution of habitual *be* in AAVE by showing that – notwithstanding the differences – there are important syntactic parallels between agentive *be* in informal American English and habitual *be* in AAVE: both are uninflected; both require do-support in the presence of negation; neither can appear to the left of negation or the left of the subject; both can occur without a preceding auxiliary; both appear to require overt subjects in finite clauses. Although Collins acknowledges Rickford's (1986) decreolization theory of habitual *be* in AAVE, he maintains that superstrate agentive *be* plausibly played a part in the historical development: it entered the emerging contact language via adult learners who encountered it in the superstrate dialects, and who subsequently adapted the syntax of agentive *be* to express habitual aspect.
– Fereirra provides a diachronic analysis of plural /s/ deletion in Afro-Hispanic and Afro-Portuguese texts (15th to 19th centuries), and finds double plural marking in the earlier texts (i.e. on the two elements of the noun phrase). This pattern was initially shared by the two varieties but diverged later with Afro-Portuguese texts marking plural only on the first element, and Afro-Hispanic texts marking it on the second element, thus showing interesting parallels with plural marking patterns in established Ibero-Romance creoles and non-standard varieties of Brazilian Portuguese and Carribean Spanish.
– Deuber looks at 'educated' Nigerian Pidgin (rather than basilectal varieties) and finds that the continuing language contact situation in Nigeria with the lexifier (English) and Nigerian Pidgin has yielded a situation more reminiscent of code-switching than what we would expect in a typical creole continuum: in those contexts where English forms occur in the speech of educated Nigerians they are fully representational of the English grammatical system, largely confined to discourse chunks or idiomatic expressions, and clearly separated from Nigerian Pidgin forms. Unlike in the Caribbean mesolects which have been documented in the literature, there are no intermediate forms, and basilectal structures "predominate even in the speech of the most educated users of the language".

- Kramer discusses serial verb constructions in Saramaccan as a stylistic feature which remains optional and transferred relatively late from the substrate. However, variation patterns shifted from the late 18th century onwards when there was an increase in the use of serial verb constructions, possibly as a marker of linguistic and cultural identity.
- Hackert considers intra-speaker variation in Bahamian Creole English based on discourse type or style. She distinguishes four discourse types in her interview data: 'chat' mode (non-narrative interview speech), narratives of personal experience (cf. Labov & Waletzky 1967), folktales and generic narratives. Looking at past inflection as a linguistic variable, she finds that narratives foster the absence of past inflection when compared to the 'chat' mode, with folktales and generic narratives showing even less inflection than narratives of personal experience. Hackert shows that in narratives of personal experience zero past marking occurs primarily in the 'complicating action' sequence of the narratives, i.e. in a functionally and structurally distinct section of the narrative "where they usually instantiate the PERFECTIVE (PFV), an aspectual category which is often unmarked crosslinguistically". The even higher percentage of zero past marking in folktales and generic narratives is explained (a) by the timeless nature of the narrative and the fact that these stories are generally "sequenced and thus temporally disambiguated" (in the case of folktales), and (b) the observation that they describe habitual situations where overt tense marking is not necessary (and frequently absent cross-linguistically; in the case of generic narratives). Hackert's analysis has important implications for data collection and fieldwork which often rely on the sociolinguistic interview, a method which, via the 'danger of death' question, favours the elicitation of narratives of personal experience, while folktales and generic narratives are not always included. As a result analyses of past inflection often show higher rates of past inflection than would have been the case if other narrative discourse types had been included.
- Parkvall takes issues with the idea of 'basilectization', i.e. the hypothesis that prolonged contact between creole and superstrate does not lead to decreolization, but that creoles have diachronically distanced themselves from the superstrate. Parkvall bases his analysis on Fattier's dialect atlas of Haitian which records the extensive variation that exists in the language, and considers both phonological (front rounded vowels, post-vocalic /r/), morpho-phonological (agglutination of determiner and noun), and morphological variables. He finds that those dialects which show more 'French' features have been influenced by modern French (which is in line with the idea of decreolization after the genesis of the contact variety), whereas dialects which are less French-like (i.e. more basilectal) only exhibit vestiges of archaic French influence (17th century features), and appear not to have been influenced by modern French. It can thus be hypothesized that their basilectal nature has historical depth and precedes contact with (modern) French. The variation patterns of whole-syllable agglutination – also a strongly basilectal feature – support this interpretation: whole-syllable agglutination is over-represented

in the basic lexicon, and almost absent from the peripheral lexicon. If Haitian was originally more like French and basilectalization a later development, then "we would expect the core to be more like French than the periphery".

– Wexler upholds that language contact yielded relexification in the case of Yiddish which he argues is a relexified West Slavic language. Moreover, this author raises the practical issue of the need to develop a dictionary model for relexified languages which would include information on the details of the relexification process, i.e. "how the substratal language determines the selection and calibration of superstratal lexicon". Wexler outlines such a lexicon model and discusses its importance not only for Yiddish but also for other contact languages such as the Caribbean Creoles.

The papers also emphasize the importance of social and historical context in understanding and explaining the outcomes of language contact. Winford, for example, looks at the enduring debate over the roles of substrate and superstrate languages in pidgin/creole genesis and shows that more exposure to the target language increases its influence on the emerging system (e.g. Haiti), while less access leads to more substrate influence (e.g. Suriname). Sanchez' paper shows how historical and social factors continue to influence the evolution of contact languages through time. She argues that although the morpheme *-ndo* is attested from early on in the history of Papiamentu, its use was initially restricted to the gerundive function and *-ndo* with progressive function was only attested from the 1930s onwards. While gerundive *-ndo* is based on Spanish, progressive *-ndo* was triggered by the increased influence of English, following the opening of a refinery and growing immigration to the islands in the 20th century. Age and social prestige are important social factors in structuring the use of progressive *-ndo*: younger speakers of upper or upper middle class families (as well as 'old' families) were more likely to use progressive *-ndo* than older speakers from less prestigious social backgrounds.

The publication of this volume would not have been possible without the trust Jacques Arends and John Singler have put into us as editors, and their encouragement and support when the manuscript preparation took longer than anticipated. We would like to dedicate this volume to the memory of Jacques Arends who died on 16 August 2005. His work on the early history of the creoles of Suriname – which also feature so strongly in this collection – has significantly shaped the development of the discipline, and has opened our eyes to the importance of demographic data and careful sociohistorical analysis. We miss him.

References

Bickerton, D. (1981). *Roots of Language.* Ann Arbor MI: Karoma.

Labov, W. & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the Verbal and Visual Arts: Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society* (pp. 12–44). Seattle WA: University of Washington Press.

Lefebvre, C. (1998). *Creole Genesis and the Acquisition of Grammar*. Cambridge: CUP.

McWhorter, J. (1998). Identifying the creole prototype: Vindicating a typological class. *Language, 74*, 788–818.

Rickford, J. R. (1986). Social contact and linguistic diffusion: Hiberno-English and New World Black English. *Language, 62*, 245–289.

PART I

Structure

The phonetics of tone in Saramaccan[1]

Jeff Good
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

This paper presents the results of a preliminary investigation of the phonetics of tone in Saramaccan, an Atlantic creole spoken in Suriname. Two particular aspects of Saramaccan tonology are focused on (i) the ways in which a phonological split between words marked for pitch accent and words marked for lexical tone are manifested phonetically and (ii) the phonetic properties of a phonological process of high-tone plateauing found in the language. The results of the study indicate that, on a phonetic level, surface tones deriving from phonological pitch accent are not distinguished from true lexical tones, and they also establish the presence of a super-high tone found only ideophones. In addition, the study mostly verifies existing descriptions of the plateauing process.

0. Introduction

This paper presents the results of a preliminary investigation of the phonetics of tone in Saramaccan, an Atlantic creole spoken in Suriname. Saramaccan has traditionally been described as exhibiting a lexical contrast between high tones and low tones, as well as allowing tone bearing units (TBU's) to be lexically unspecified for tone, but surfacing with either high or low tone, depending on the phonological environment in which they appear (Voorhoeve 1961; Rountree 1972).

The tone system of Saramaccan is particularly well developed for a creole – this fact alone makes an examination of the phonetics of the language's tones of potential value. In addition, there are at least two other interesting aspects of Saramaccan tonology. The first is that, as reported extensively in Good (2004), a range of phonological evidence indicates that the Saramaccan lexicon exhibits a prosodic split wherein most words are marked for lexically-contrastive pitch accent but an important minority of words are marked for true tone – no other language has been explicitly

1. I would like to thank Larry Hyman, John McWhorter, and Irina Galichenko for their extensive contributions to this paper. Thanks are also due to Sharon Inkelas, Ian Maddieson, Marvin Kramer, Catherine Rountree, audience members at the SPCL conference in Atlanta in January 2003, and two anonymous reviewers. Finally, I would like to give special thanks to my principal consultants.

reported as showing such a split.[2] The second area of interest is a process of high-tone plateauing found in phonological phrases in the language which has been the topic of both descriptive and theoretical studies (Voorhoeve 1961; Rountree 1972; Devonish 1989: 48–55; Ham 1999; Good 2003).

Until now no instrumental phonetic data has been available for Saramaccan tone. While the data to be discussed here falls short of a complete survey, it is hoped that it will be of value both to the study of Saramaccan specifically and to comparative studies of the prosodic systems of Atlantic creoles (and other contact languages) more generally.

The structure of this paper is as follows: Section 1 summarizes some basic descriptive features of tone in Saramaccan. Section 2 briefly examines some of the evidence for a pitch accent/tone split in the Saramaccan lexicon. Section 3 presents the results of a phonetic study of the pitch of words uttered in isolation, and Section 4 presents some of the results of an examination of pitch in phrases. Finally, Section 5 offers a brief conclusion.

1. The three-way tonal lexical contrast in Saramaccan

There is a three-way lexical contrast in Saramaccan among tone bearing units (TBU's) marked for high tone, low tone, and those which are unspecified for tone. By default, TBU's unspecified for tone surface as low, but, in certain well-defined phonosyntactic environments, they surface as high.

The primary evidence for a three-way contrast, as opposed to a two-way contrast, comes from data involving a process of high-tone plateauing which only affects TBU's unspecified for tone. Relevant examples are given in (1).[3]

---

2. Two languages are explicitly reported as exhibiting phenomena which bear some similarity to the Saramaccan situation. One of these is Papiamentu, another Atlantic creole, which has been reported to make use of both contrastive pitch accent and stress (Kouwenberg & Murray 1994; Rivera-Castillo 1998; Remijsen 2002: 43; Rivera-Castillo & Pickering 2004). Another language, Ma'ya, of the Austronesian family, has been reported as having both contrastive tone and contrastive stress (Remijsen 2001; Remijsen 2002: 39–68). While neither of these languages show *split* lexicons like Saramaccan, they represent cases where a language makes use of two different prosodic systems to make lexical contrasts. Devonish's (2002) analyses of Guyanese Creole (2002: 82–119) and, perhaps, Krio (2002: 147) indicate they may exhibit something similar to what will be reported here for Saramaccan. He does not explicitly describe them as having split lexicons, but, clearly, they would be worthwhile to examine in trying to find further examples of this type of phenomenon.

3. Throughout this paper, the following conventions will be maintained: an acute accent (´) will be used to mark high tone and a grave (`) will be used to mark low tone. Surface forms will be completely tone marked, reflecting their actual pronunciation. Underlying forms, however, will only show the tone marking which is taken to result from their lexical specification. Ortho-

(1) a. *taánga* → *tàángà* 'strong'
b. *dí taánga wómi* → *dí tàángá wómì*
the strong man
'the strong man'

In (1a), the word *taánga* 'strong' is seen in its citation form outside of a phrase. As indicated in the example, the form of the word in this context is *tàángà*, where the two TBU's lexically unspecified for tone surface with low tones. In (1b), the word *taánga* appears in a noun phrase, and it surfaces with a final high tone as *tàángá*. The appearance of this high tone follows a general rule that unspecified tones are realized as high tones within a phonological phrase when flanked by high tones. Thus, this yields a type of tonal plateauing, creating one flat series of high tones within phonological phrases where otherwise there might have been a contour. The particular phonological phrase involved in the plateauing seen in (1b) is one formed by a noun and the word preceding it in the noun phrase.[4] For detailed discussion on the syntactic environments constituting phonological phrases in Saramaccan, see Good (2004: 598–607).

Since, in a word like *taánga*, the surface tones on the first and last TBU are predictable based on their phonosyntactic environment, they are most straightforwardly analyzed as being unspecified for tone. A word like *taánga*, then, can be usefully contrasted with a word like *káìmà* 'alligator'. The citation form of this word also contains some high tones and low tones. However, unlike *taánga*, the low tones in *káìmà* never appear as high. The examples in (2) show how *káìmà* contrasts with the word *wómi* 'man', whose final TBU (like that of *taánga*) is lexically unspecified for tone.

(2) a. *Dí káìmà kulé àlá.* → *Dí káìmà kùlé àlá.*
the alligator run there
'The alligator runs there.' (Rountree 1972: 316)
b. *Dí wómi kulé àlá.* → *Dí wómí kúlé àlá.*
the man run there
'The man runs there.'

The last word of a subject noun phrase forms a phonological phrase with a following verb. Thus, the unspecified final TBU of *wómi* and the unspecified initial TBU of *kulé* 'run' both surface with high tones, as a result of plateauing, in (2b). However, in the same basic environment, the final two TBU's of *káìmà* surface as low reflecting the fact that they are lexically specified for their low tones and, therefore, are never affected by plateauing. In addition, the final low-tone TBU's of *káìmà* block the possibility that

graphic *ng* is a velar nasal and 'coda' *n*'s indicate nasalization on the preceding vowel. All uncited data comes from consultant work.

4. The data in (1b) also shows that an adjective and a preceding article do not form a phonological phrase – hence, the initial TBU of *taánga* surfaces as low in (1b) even though it is flanked by high tones.

high-tone plateauing could affect the word *kulé*. Thus, unlike in the sentence in (2b), in (2a), *kulé* surfaces with an initial low-tone TBU.

In the next section, the distribution of different tone patterns in Saramaccan words will be discussed, with a focus on evidence that the language's lexicon is "split" into an accentual part and a tonal part.

2. The "split" lexicon of Saramaccan

2.1 Commonly attested tonal patterns

In this section, I will discuss some of the evidence for a lexical split in Saramaccan between a class of words marked for accent and a class marked for tone. Aspects of the analysis I will provide are anticipated by Devonish (1989: 48–55) and Devonish (2002: 120–134), in particular the idea that the majority of words in Saramaccan are marked for pitch accent. While he does not specifically interpret the Saramaccan lexicon as being 'split', Devonish's analyses and the one given here are largely compatible.[5]

The first step in recognizing the split in the Saramaccan lexicon lies in the observation that, while there is fairly good evidence for an underlying distinction between three types of TBU's in Saramaccan – high tone, low tone, and unspecified for tone – the observable contour patterns on words containing TBU's unspecified for tone are quite limited. All the common patterns are exemplified in Table 1.[6]

Table 1. Common patterns of words with unspecified TBU's

WORD	TONES	GLOSS
foló	ØH	'flower'
náki	HØ	'hit'
sikífi	ØHØ	'write'
mɔ́kisá	ØØH	'screen, sift'
hákísi	HHØ	'ask'
afokáti	ØØHØ	'lawyer'
minísíti	ØHHØ	'minister'
alukutú	ØØØH	'soursop (fruit)'

5. I believe some of the differences between the analysis here and that of Devonish (2002: 120–134) lie in the fact that his analysis relied on published sources, where the tones of words are often inconsistently transcribed, while I made use of data from published sources and also had the opportunity to collect data from consultants to verify those sources.

6. The tonal contours in Table 1 are solely meant to serve as a schematic way of representing a particular, attested pattern. They are not intended to be interpreted as underlying representations.

Words containing TBU's unspecified for tone comprise the majority of the Saramaccan lexicon (Rountree 1972: 316). Accurate figures are not available, but probably around ninety percent of monomorphemic words of this type.

One of the most noteworthy restrictions on words with TBU's unspecified for tone, which comes out clearly in Table 1, is that none of the common types contain low tones.[7] Another restriction is that most words of this type only surface with one high-tone TBU in their citation form. As can be seen in the table, in some cases, the words do have multiple high-tone TBU's, but, when this is the case, the high-tone TBU's must be adjacent and are always the antepenultimate and penultimate TBU's of the word. As discussed in detail in Good (2004), the distribution of high tones in these words has the characteristics of a pitch accent system instead of a true tone system.[8] By *pitch accent*, I mean a type of lexical marking wherein one position in a word is specified for abstract phonological prominence and the primary phonetic realization of prominence is via a consistent tonal pattern – in Saramaccan, this tone pattern would be a simple high tone. Pitch-accent systems can, on the one hand, be opposed to stress-accent systems, like that of English, where the primary realization of prominence is through non-pitch cues like the amplitude, duration, and vowel quality of a particular syllable.[9] And, they can, on the other hand, be opposed to tone systems, where pitch is assigned to words lexically and is not a reflex of phonological prominence at all.

To understand how a pitch accent analysis of words with TBU's unspecified for tone would work, it is first useful to group the major tone patterns seen in Table 1 with respect to CV structures, which is done in (3).

(3)	**2-σ words:**	CV́CV	CVCV́	
	3-σ words:	CV́CV́CV	CVCV́CV	CVCVCV́
	4-σ words:	CVCV́CV́CV	CVCVCV́CV	CVCVCVCV́

In (4) the words from Table 1 exemplifying the patterns in (3) are given.

---

7. I have identified one monomorphemic word containing both a lexical low tone and a TBU unspecified for tone, *anákìtá* 'biting ant', which has the tonal form ØHLH. Voorhoeve (1961: 154) identifies about ten words (out of a sample of 1500 words) which follow a similar pattern – an initial TBU unspecified for tone with lexical low tone in some other position in the word. All but one of the words he gives begin with *a* like *anákìtá*, and the one exception to this, *obílògbén* 'a type of snake', also begins with a vowel.

8. Good (2004) contains a number of arguments for the position that words with TBU's unspecified for tone should be analyzed as being part of a pitch accent system wherein other words in Saramaccan are truly tonal. It will not be possible to go through all of those arguments here.

9. The distinction between these two ideal types of accent systems can sometimes be obscured in stress-accent languages, like English, where the realization of an intonational pitch contour is sensitive to the position of stress-accent in the words comprising an utterance. The difference between a pitch-accent language and a language like English is that, in English, there is not a consistent tonal contour associated with a stressed syllable. Rather this tonal contour can change depending on the nature of the intonational contour associated with the utterance.

(4) **2-σ words:** *náki* *foló*
3-σ words: *hákísi* *sikífi* *mɔkisá*
4-σ words: *minísíti* *afokáti* *alukutú*

The restricted possibilities for the tonal patterns in words containing TBU's unspecified for tone allows for a 'one-mark-per-word' analysis wherein the specification of one TBU in a word for prominence allows us to predict where the high tone or high tones will appear. The location of the necessary prominence 'mark', with respect to the CV schematization seen in (3), is given in (5), where an asterisk above a vowel indicates abstract phonological prominence.

(5) **2-σ words:** C$\overset{*}{\text{V}}$CV CVC$\overset{*}{\text{V}}$
3-σ words: C$\overset{*}{\text{V}}$CVCV CVC$\overset{*}{\text{V}}$CV CVCVC$\overset{*}{\text{V}}$
4-σ words: CVC$\overset{*}{\text{V}}$CVCV CVCVC$\overset{*}{\text{V}}$CV CVCVCVC$\overset{*}{\text{V}}$

The high-tone pattern on words with TBU's unspecified for tone is completely predictable from the position of the abstract prominence marks given in the schematized word structures in (5). In some cases, the relationship between prominence and high-tone marking is trivial – a vowel marked for prominence simply surfaces with a high tone. However, the relationship is not always so simple – in words with a prominence mark in antepenultimate position a high tone is realized on both the antepenultimate and penultimate TBU. If there were a class of words showing contours like (Ø)HØØ in Saramaccan, the analysis of words with (Ø)HHØ as showing the reflex of antepenultimate prominence would be problematic. However, a conspicuous lack of (Ø)HØØ contours in the language makes such an analysis straightforward.

The fact that the position of the high tone or high tones in words containing TBU's unspecified for tone can be predicted on the basis of one abstract lexical 'mark' means that they belong to a system which marks syntagmatic contrast – that is the locus of the contrast is the relationship different positions within the word have with respect to each other. As discussed by Hyman (1978:7), this is a central characteristic of an accent system.

If the Saramaccan lexicon consisted only of words of the type just discussed, the language could be described simply as a pitch accent language. However, not all words belong to one of the types exemplified in Table 1. In (2a), for example, we saw the word *káìmà* 'alligator', which is fully marked for tone. Furthermore, there are cases in Saramaccan where one clearly finds pitch being used to mark paradigmatic contrast – that is, two words lexically contrast solely on the basis of the pitch used to mark one of the TBU's in each word. Hyman (1978:7) gives the use of pitch to mark such contrast as a central characteristic of a tone system. One example of a pitch-based paradigmatic contrast in Saramaccan can be found in the minimal pair *fà* 'fun' and *fá* 'manner', two words which differ phonologically only in the pitch assigned to their one TBU.

Table 2. Emphatic and non-emphatic pronouns in Saramaccan

	NON-EMPHATIC		EMPHATIC	
PER	SG	PL	SG	PL
1st	*mì*	*ù*	*mí*	*ú*
2nd	*ì*	*ùn*	*í*	*ún*
3rd	*à*	*dè*	*hén*	*dé*

However, the clearest case of tone being used to mark paradigmatic contrast comes from the opposition between 'non-emphatic' and 'emphatic' pronominal forms in the language.[10] These are given in Table 2, as reported in Voorhoeve (1961: 161).[11] As can be seen in Table 2, for five of the six pronominal forms, the only way the emphatic form is distinguished from the non-emphatic form is via a change in tonal marking.

In addition to showing such paradigmatic contrasts, words fully marked for tone further differ from words containing TBU's unspecified for tone insofar as they exhibit a wide range of possible tone patterns, including patterns consisting of multiple high-tone TBU's, multiple low-tone TBU's, and patterns where both high-tone and low-tone TBU's are found within a single word. Examples are given in Table 3.

There is good evidence, then, that while most words in Saramaccan may be part of a pitch accent system, there is a separate class of words which appear to be part of a true tone system, showing both paradigmatic pitch contrasts and a wide range of tonal contours. In the next section, I will briefly comment on the likely historical origins of this split in the Saramaccan lexicon.

2.2 On the origins of the split

The most likely account of the origins of the split lexicon in Saramaccan is that the language exhibits a logically possible (but otherwise unattested) contact phenomenon between European accent languages and African tone languages. Rather than 'levelling' its lexicon towards an African type or a European type, Saramaccan appears to have

10. The distribution of these two sets of pronouns is complicated and not simply conditioned by the pragmatic parameter of non-emphatic versus emphatic. For example, object pronouns, except for the third plural, are drawn from the 'emphatic' paradigm (Voorhoeve 1961: 161).

11. A possible alternative way of analyzing the 'tonal' opposition of the pronominal forms in Table 2 would be to say that the non-emphatic pronouns are underlyingly unaccented while the emphatic ones are underlyingly accented. My primary reason for adopting the tonal analysis given here is that, as pointed out by Voorhoeve (1961: 161), both paradigms of pronouns are subject to reductions, indicating that neither is truly stressed and, as discussed in Good (2004: 588–592), there is otherwise a close correlation between a pitch accent high tone and stress, which implies that, if the pronouns of the emphatic paradigm were marked for pitch accent, they should also be stressed and, therefore, not subject to reduction.

Table 3. Tone patterns for words fully specified for tone

	WORD	TONES	GLOSS
High tones only	*hɛ́n*	H	'he'
	sósó	HH	'only'
	búúú	HHH	'ideophone for covering'
Low tones only	*bà*	L	'carry (mass noun)'
	bɔ̀sɔ̀	LL	'loosen'
	lɛ̀gɛ̀dɛ̀	LLL	'lie'
High and low tones	*àkí*	LH	'here'
	káìmà	HLL	'alligator'
	tótómbòtí	HHLH	'woodpecker'
	séségùùsɛ́	HHLLH	'kind of fish'

Table 4. Some Saramaccan words of European origin

SARAMACCAN	GLOSS	ORIGIN	
náki	'hit'	< English	*knock*
kulé	'run'	< Portuguese	*correr*
sitónu	'stone'	< English	*stone*
síkísi	'six'	< English	*six*
wólúku	'cloud'	< Dutch	*wolk*
minísíti	'minister'	< Dutch	*minister*
amɛɛkán	'American'	< Dutch	*Amerikaan*

maintained two parallel prosodic systems, one, roughly speaking, with "European" characteristics and another with 'African' characteristics.[12]

Some evidence for this historical scenario comes from the fact that, as pointed out by Ham (1999: 55), in transferred words of European origin, the high-tone TBU in Saramaccan tends to correspond to the nucleus of the stressed syllable in the relevant European language. Additionally, in words with two adjacent high-tone TBU's across two syllables, the general pattern is that the initial TBU corresponds to a TBU in the stressed syllable of the source language and the second high-tone TBU is a historically epenthetic vowel. Relevant examples are given in Table 4.

There are exceptions to these generalizations. For example, following the pattern exemplified in Table 4, the word *àkí* 'here', from Portuguese *aqui*, would be expected to have been transferred into Saramaccan as *akí* with its first TBU unspecified for tone.

12. As discussed by Bruyn (2002: 165–167), Amerindian elements can also be identified in Surinamese creoles. It is not clear what general principles, if any, govern their prosodic phonology in Saramaccan. For example, the Saramaccan word *amáka* 'hammock', identified by Bruyn (2002: 166) as being of Arawakan origin, is apparently marked for pitch accent, while the word *káìmà* 'alligator', which (controversially) is of Carib origin, according to the Oxford English Dictionary, is apparently marked for true tone.

Table 5. Some Saramaccan words of Kongo origin

SARAMACCAN	GLOSS	ORIGIN	
pùkùsù	'bat'	< Kongo	*lu-mpukusu*
bàndjà	'side'	< Kongo	*mbaansya*
mbàlù	'(wood) chips'	< Kongo	*mbalu*
màtùtù	'small rat'	< Kongo	*ma-tutu* (pl.)

However, the word *àkí* is instead fully marked for tone. Nevertheless, the fact that there is a strong correlation between accent in European languages and pitch-accent in Saramaccan strongly indicates that accent entered the language via transfer from European languages.

Though the case is more tenuous because of the lack of sufficient data, there is some indication that words fully marked for tone tend to be of African origin, giving us evidence that the existence of the tonal part of the Saramaccan lexicon has its roots in the transfer of African prosodic systems. Daeleman (1972:2), for example, notes a correlation between words fully marked for low tone in Saramaccan with comparable words in Kongo. Examples, taken from Daeleman (1972), are given in Table 5.

As Daeleman (1972:5) points out, there is not always complete correspondence between Kongo tones and Saramaccan tones. So, the story is more complicated than simply stating that tonal words in Saramaccan have their tones as a result of direct transfer from an African language. Nevertheless, the data in Tables 4 and 5 gives us initial evidence, at least, that the split between tonal and accentual words in the Saramaccan lexicon is, broadly speaking, the result of transfer of both African tonal systems and European accentual systems into the language without levelling of the prosodic structure of words in favor of one system over the other.

3. A preliminary phonetic examination of the split lexicon

3.1 Theoretical background

Given the apparent split *phonology* of the Saramaccan lexicon, an important question is whether or not this split has any detectable phonetic reflex. This issue is of more than purely descriptive interest given the assumptions of works like Ladd (1996) (and related references cited therein) which assume that, except for the level of constituency where it is assigned, there is no fundamental phonological difference between lexically-assigned pitch (i.e. tones) and pitch assigned as a result of pitch-accent or intonation.

Under such a view, both tonal languages and intonational languages are understood to be making use of 'tones' in the assignment of pitch contours. In a tone language, tones are assigned to particular positions in words within the lexicon, while, in an intonational language, tones are assigned to larger constituents, like sentences. In a pitch accent language, tones are realized in particular positions in a word, like

in a tone language. However, tone itself is not assigned lexically – only prominence is lexical, with tone assignment being a predictable reflex of prominence.

Saramaccan is, at present, the only language explicitly analyzed as having a split tonal/accentual phonology along the lines described here. It, therefore, serves as an interesting test case for claims like those found in Ladd (1996). If those claims are correct, lexically-specified high tones should be indistinguishable, in terms of F_0 targets, from high tones realized as the result of pitch accent, and lexically-specified low tones should be similarly indistinguishable from 'default' low tones (that is, the low tones appearing on TBU's unspecified for tone when they have not undergone high-tone plateauing of the sort discussed in Section 1).

3.2 Data collection

To determine the overall phonetic properties of Saramaccan tone, with a focus on possible differences between tones assigned lexically versus those assigned accentually, a number of words were each elicited twice in isolation from a male native speaker. These words were chosen so as to exemplify the various tonal combinations found in Saramaccan as first classified by Rountree (1972: 314–18).

The recordings of these words were analyzed in Praat and measurements of F_0 were made for each TBU – except for ideophones, where a string of adjacent identical vowels was treated as one TBU.[13] Ideophones are a special lexical class of words in Saramaccan used for emphasis and not typically integrated into the syntax of sentences but, instead, used exclamatively. They differ phonologically from other words in Saramaccan in that they generally must be specified as having all high-tone TBU's or low-tone TBU's (in addition to having other specific phonological characteristics, like a strong tendency towards total vowel harmony).

The words in Table 6 were the words used in this study analyzed as being part of the tonal part of the Saramaccan lexicon, and the words in Table 7 were the words used in this study analyzed as being part of the pitch-accent part of the Saramaccan lexicon.

The pitch measurement for each TBU was measured at the highest F_0 value for surfacing high-tone TBU's and the lowest F_0 value for surfacing low-tone TBU's. (This method of measurement is consistent with Ladd (1996) and was explicitly adopted by Liberman (1993).) Table 8 summarizes the results of this data collection. The mean F_0 is given for three different classes of high tones and three different classes of low tones. For high tones these classes were: high tones in ideophones, lexically-specified high tones, and pitch-accent high tones. For low tones, these classes were: low tones in ideophones, lexically-specified low tones, and default low tones.[14] In addition to giving

13. Praat is a free speech-analysis program. Detailed information about it can be found at http://www.praat.org/.

14. At the beginning of the study, ideophones were not intended to be treated as a separate class for pitch measurement. However, a cursory examination of the mean F_0 values for high

Table 6. Words fully marked for tone elicited for this study

TONE TYPE	WORD	GLOSS
High	*búúú*	'ideophone for covering'
	hɛ́n	'he.EMPH'
	fáán	'ideophone for white'
	kúlúlúú	'ideophone for straight'
Low	*bà*	'carry (liquid)'
	bɛ̀	'red'
	bɔ̀sɔ̀	'loose'
	lègèdè	'lie'
	kùnàkùnà	'old, worn out'
	pètèpètrè	'ideophone for salve-like'
Mixed	*àkí*	'here'
	kâìmà	'alligator'
	sɛ́sɛ́gùùsɛ́	'type of fish'
	tótómbòtí	'woodpecker'

Table 7. Pitch-accent words elicited for this study

WORD	GLOSS
amɛɛkán	'American'
éi	'if'
foló	'flower'
folóísi	'crowd'
maaní	'screen, sift'
náki	'hit'
ingíísi	'English'
sóso	'only'
sukuáti	'chocolate drink'
taánga	'strong'

the mean F_0 value (in Hertz) for each type of TBU, the table also gives the standard deviation and range of F_0 as well as the number of TBU's of each type measured in the study.

An informal examination of the table reveals first and foremost that the high tones in ideophones were realized with much higher pitch than the other types of high tones, which had relatively similar pitch values to each other. In addition, all three types of low tones had roughly similar pitch values.

Not surprisingly given the very high pitch range of the high-tone ideophones, the differences among the three types of high tones were found to be highly statistically significant (a one way ANOVA test over the three types of high tones gave a p-value of

tones in ideophones made it immediately clear that they strongly diverged from the other high tones, which is why ideophone tones were separated from the other tone types here.

Table 8. Pitch figures for different phonological sources of tone in Saramaccan

	MEAN F_0	ST.DEV.	RANGE	NO.
High-tone ideophones	162	3.8	11	10
Lexical high tones	131	7.3	24	18
Accent high tones	127	5.1	18	24
Low-tone ideophones	107	3.8	12	8
Lexical low tones	104	4.9	18	34
Low-tone defaults	106	7.0	35	36

effectively zero).[15] Lexically-specified high tones, however, were not found to be significantly different from high tones resulting from pitch accent (two-tailed t-test, p-value 0.068).[16] Finally, the three types of low tones were also not found to be significantly different from each other (one way ANOVA, p-value 0.35).

Both informally and statistically, then, these measurements indicate several things. First, the high tones in ideophones are clearly different from other high tones – in fact, the data points to an analysis where ideophones should be considered to be marked with something like a 'super-high' tone. These measurements also indicate that, from a phonetic standpoint, the F_0 maxima of lexical high tones and pitch accent high tones are the same, consistent with Ladd's (1996) hypotheses about the relationship between lexical tones and other types of tone. In addition, they indicate that lexical low tones and default low tones have the same F_0 minima, also consistent with Ladd's (1996) hypotheses. Finally, unlike high-tone ideophones, low-tone ideophones appear to make use of the same F_0 target as other low tones.

The results seen here, then, favor the basic hypotheses espoused by Ladd (1996). Despite their different phonological sources, the F_0 maxima and minima of high tones and low tones respectively in words marked for pitch accent, as opposed to tone, appear to be the same. Ideophones complicate the picture somewhat. In the case of high-tone ideophones, the results suggest a third phonological tone in Saramaccan – the super-high tone. This fact is not particularly striking since, as mentioned above, ideophones differ phonologically in a range of ways from non-ideophones in the language. In the case of low-tone ideophones, there is no such complication – they appear to have the same F_0 target as the two other types of low tones.

15. All statistical calculations were done using Microsoft Excel.

16. This result is much closer to being significant than the differences between low tones, indicating that the F_0 maxima for lexical high tones versus pitch accent high tones should be examined closely in future work.

4. The phonetics of high-tone plateauing

4.1 Introduction

As discussed above in Section 1, one of the areas of Saramaccan tonology which has been of interest to previous researchers is a process of tonal plateauing wherein TBU's unspecified for tone are realized with high tones, instead of default low tones, when flanked by high tones within a phonological phrase. Rountree (1972) and Good (2004) offer fairly detailed discussion of where high-tone plateauing occurs. Here, I will present limited data on the phonetics of two environments where plateauing is found: between a noun and a preceding adjective and between a subject and a following verb. The data will be in the form of pitch traces of particular phrases. These will primarily be examined to determine whether or not they are consistent with the reported descriptions, but, where relevant, other points of potential interest will be discussed.

4.2 Plateauing between a noun and preceding adjective

As seen above in (1), an adjective and following noun form a high-tone plateauing environment. The examples in (1) are repeated below in (6).

(6) a. *taánga* → *tàángà* 'strong'
b. *dí taánga wómi* → *dí tàángá wómì*
the strong man
'the strong man'

In (6a) the citation form for *taánga* 'strong' is given where default low tones appear on its first and last TBU's. In (6b) the final TBU of *taánga* is realized with a high tone since it is flanked by two high tones and is part of a phonological phrase consisting of a noun and a preceding adjective. Pitch traces for (6a) and (6b) are given in Figures 1 and 2 respectively.

The pitch traces in Figure 1 and Figure 2 are both consistent with the reported descriptions. In Figure 1, the contour starts relatively low, reaches a maximum over the second TBU of *taánga*, and then falls over the last TBU, as expected given the transcription. In Figure 2, the sequence *tàángá wómì* roughly matches the pattern indicated in the transcription. Of particular interest is the lack of any drop over the last TBU of *taánga* – this is the TBU described as being realized as high as the result of plateauing.

A noteworthy property of the pitch trace in Figure 2, however, is that it is not completely level between the second TBU of *taánga* and the first TBU of *wómi* 'man'. Rather, the pitch gradually moves upward, reaching its highest point above the first TBU of *wómi*. This gradual movement, in fact, calls into question the accuracy of the transcription *dí tàángá wómì* given how slight the pitch difference is between the first TBU of *taánga* and the two subsequent TBU's. Transcribing adjective-noun combina-

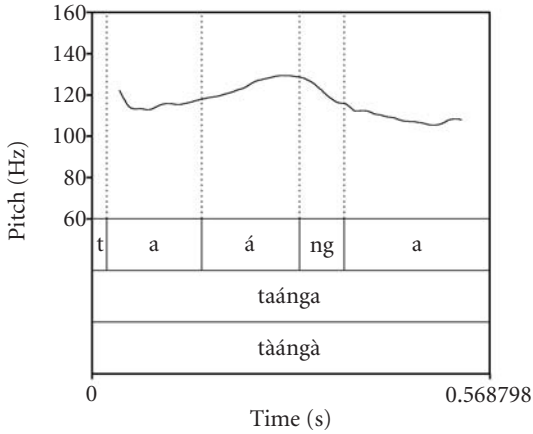


Figure 1. Pitch trace and segmentation of *tàángà* 'strong'

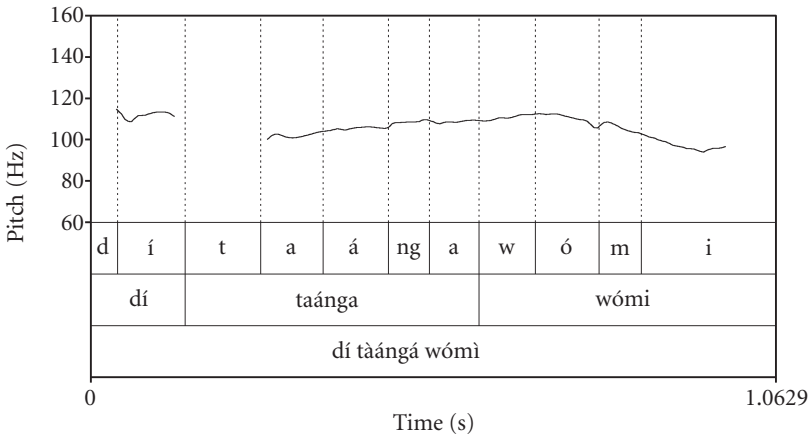


Figure 2. Pitch trace and segmentation of *dí tàángá wómì* 'the strong man'

tions as exhibiting plateauing is not new to this work but, in fact, has precedents in Voorhoeve (1961: 159) and Rountree (1972: 321).

This discrepancy between the pitch trace and these transcriptions is suggestive of a system wherein, within the phonological phrase *taángá wómi*, a second phenomenon, in addition to plateauing, can be observed – specifically, the head noun of the phrase may be receiving primary accent while the adjective only has secondary accent. The phonetic reflex of such an accent pattern would be the higher pitch over *wómi* as compared to *taánga* as well as the relative lack of differentiation between the pitch levels of the first TBU of *taánga* and the two subsequent TBU's.

If the two words in the phrase are being marked for different levels of accent, Saramaccan would be showing behavior attested in other languages which mark their

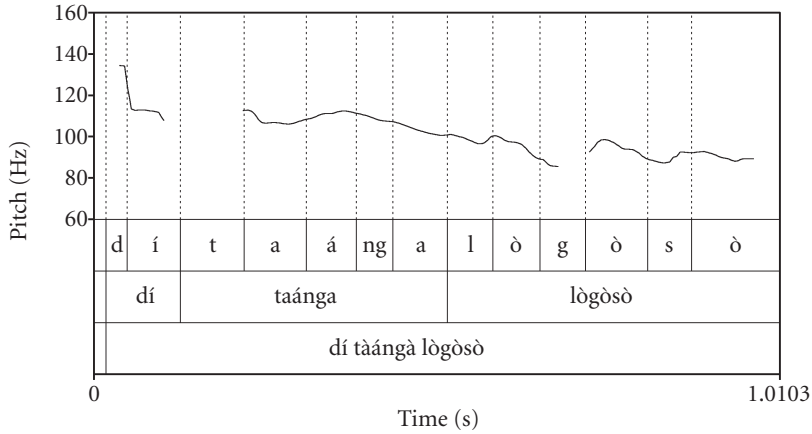


Figure 3. Pitch trace and segmentation of *dí tàángà lògòsò* 'the strong turtle'

words for accent. For example, in English, a stress-accent language, the phrase *white house* receives primary stress on the noun and the adjective is only secondarily stressed, as is clear from the opposition between that phrase and the compound *White House*. Of course, evidence from one pitch trace can only be taken as suggestive of the idea that Saramaccan shows primary and secondary accent phenomena in phrases. Nevertheless, Figure 2 indicates that this is an area worthy of further study, especially in light of arguments in Good (2004: 602–607) which suggest, on purely phonological grounds, that the phrasal phonology of Saramaccan more closely resembles that of an accentual language than a tonal one, despite the presence of a class of tonal words in its lexicon. It would seem that phonetic data, in addition to phonological data, could have bearing on this issue.

High-tone plateauing, of the sort seen in (6b), between nouns and adjectives has been reported as being blocked when the head noun begins with a low-tone TBU. An example of such a noun phrase is given in (7). This example is syntactically parallel to (6b). As indicated in the transcription, the last TBU of *taánga* has been observed to surface with a low tone when preceding the low-tone noun *lògòsò* 'turtle'.

(7) *dí taánga lògòsò* → *dí tàángà lògòsò*
the strong turtle
'the strong turtle'

The effect indicated by the transcription in (7) can be clearly observed in the pitch trace of the phrase seen in Figure 3. In particular, the pitch trace of the word *taánga* takes on a shape which is roughly similar to the shape it has in isolation in Figure 1, and it markedly differs from its shape in Figure 2.

In this section, we have seen some data on high-tone plateauing between a noun and preceding adjective. The pitch traces verify previous descriptions of the language, and the pitch trace in Figure 2 further indicates that a worthwhile future area of re-

search would be an examination of the possibility that Saramaccan uses varying pitch levels to mark words in phrases for primary and secondary accent.

In the next section, I will discuss data regarding high-tone plateauing found between the last word of a subject noun phrase and a following verb.

4.3 Plateauing between a subject and predicate

In addition to taking place between a noun and preceding adjective, another environment where plateauing has been reported to occur is between the last word in a subject noun phrase (typically the head noun) and the verb in the following predicate. (Basic Saramaccan word order is SVO.) The sentence in (8) shows this type of plateauing. Its pitch trace is given in Figure 4. The second vowel in *bóto* is a TBU reported as being affected by the plateauing, and the pitch trace clearly reflects this.

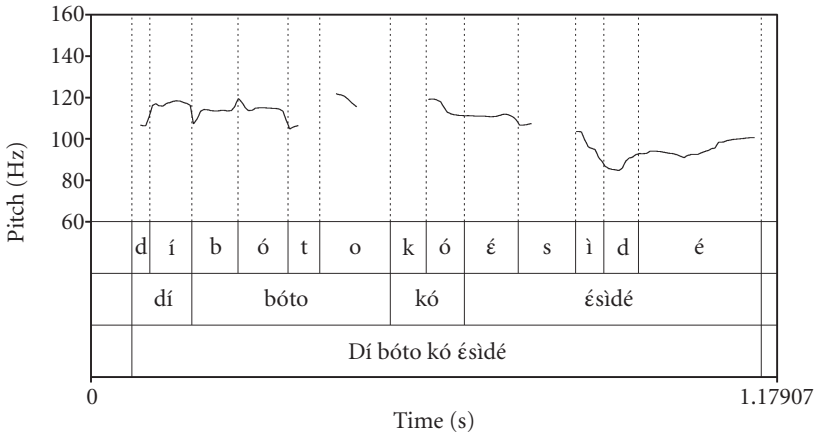


Figure 4. Pitch trace and segmentation of *dí bótó kó ɛ́sìdé* 'the boat came yesterday'

(8) *Dí bóto kó ɛ́sìdè.* → *Dí bótó kó ɛ́sìdè.*
the boat come yesterday
'The boat came yesterday.'

The sentence in (9) gives a sentence where high-tone plateauing is blocked by the presence of the low-tone noun in the subject. The pitch trace for this sentence is given in Figure 5.[17]

(9) *Dí lògòsò kulé àlá.* → *Dí lògòsò kùlé àlá.*
the turtle run there
'The turtle ran there.'

17. The final high tones of the sentences in (9) and (10) are affected by an untranscribed process of utterance-final lowering (see Rountree (1972: 309)).

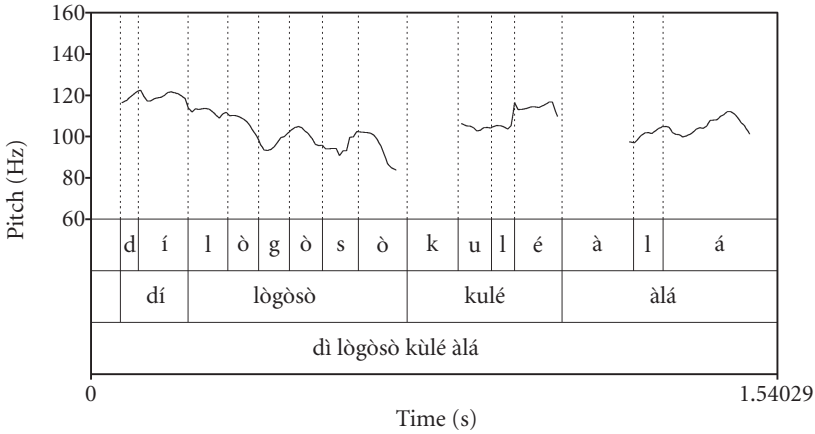


Figure 5. Pitch trace and segmentation of *dí lògòsò kùlé àlá* 'the turtle ran there'

As can be seen, the pitch trace in Figure 5 is consistent with the reported descriptions since the first TBU of *kulé* 'run', though higher than the last TBU of *lògòsò*, is relatively close to it as well as being markedly lower than the second TBU of the word, indicating plateauing has not taken place.[18]

The pitch trace for the sentence in (10), in principle, should straightforwardly contrast with the pitch trace for the sentence in (9). While the two sentences are not exactly parallel, they are very similar syntactically. Critically, they both contain the same verb *kulé*. Unlike (9), however, the subject of (10) is not a noun phrase headed by a low-tone noun. Rather, it is a high-tone emphatic pronoun *hén*. The pitch trace of (10) is given in Figure 6.[19]

(10) *Dí wómi, hén kulé dé.* → *Dí wómì, hén kúlé dé.*
the man he.EMPH run there
'The man, he runs there.' (Rountree 1972: 324)

As can be seen, the first TBU of *kulé* in Figure 6 is clearly higher than it is in Figure 5, consistent with the fact that plateauing is reported in sentence (10), but not in sentence (9). The pitch contour starts surprisingly high in the first syllable of *kulé* in Figure 6,

18. The final vowel in *lògòsò* in the token used to generate Figure 5 was partially devoiced. An artifact of this is the steep drop in pitch trace at the end of the word. Since this part of the pitch trace appears over a portion of the vowel where voicing was particularly weak, it is not as reliable an indicator of F_0 as the part of the pitch trace found over the beginning of the vowel, which is what is being compared to the pitch trace over the first syllable of *kulé* here.

19. In the particular utterance examined in Figure 6, the speaker used a reduced form of *hén* without the initial *h*, which is why this segment does not appear in the segmental tier of the transcription.

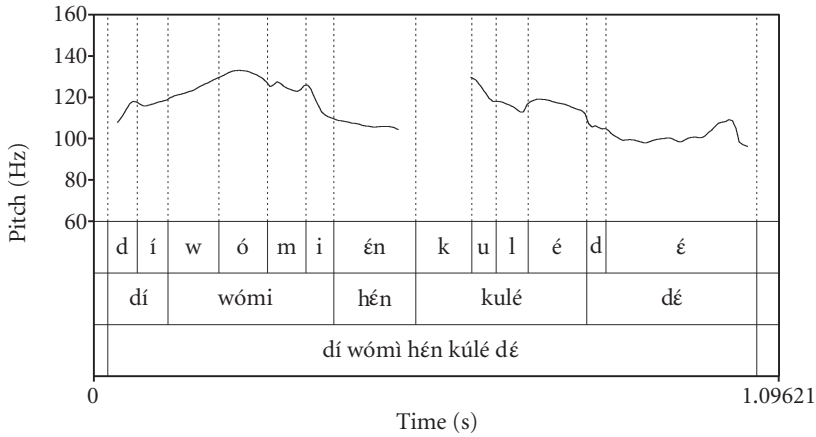


Figure 6. Pitch trace and segmentation of *dí wómì, hɛ́n kúlé dɛ̀* 'the man, he runs there'

but this might be an effect of the voiceless stop at the beginning of the word – I will come back to this issue briefly in Section 5.

There is another, less easily explained, aspect of Figure 6, however. The traditional description of the sentence in (10) would suggest that the last TBU of *wómi* would surface as low. However, it is slightly higher (and, in fact, falls into) the transcribed high tone of *hɛ́n*. This suggests some complex interaction between the pitch contour assigned to the initial appositive noun phrase and the pitch contour of the whole sentence. While not necessarily contradicting traditional descriptions, the pitch trace in Figure 6 indicates that there are unresolved issues in the study of Saramaccan pitch contours at the level of the utterance and that, while phonemicized transcriptions of each TBU as high or low might be appropriate for some levels of description, they are not appropriate for all phenomena.

5. Conclusion

This study has been intended only as a preliminary analysis and clearly further investigation is needed in order to solidify the conclusions reached here. Nevertheless, it is possible to make some tentative statements about the phonetics of Saramaccan tone at this point.

First, an examination of words in isolation indicates that the F_0 targets of tones assigned to words phonologically marked for pitch accent and words phonologically marked for lexical tone are the same. This makes the Saramaccan system consistent with ideas discussed in Ladd (1996) that the difference between tone, pitch accent, and intonation is not whether or not a language uses tone but, rather, at what level of constituency tones are associated with segmental material. This is a relatively striking result given that the hypotheses in Ladd (1996) were not designed with a split system

like Saramaccan's in mind, for the simple reason that no other language has been explicitly reported as exhibiting such a split. This preliminary study, therefore, gives a new line of support to Ladd (1996) and related work.

Another result of this study is that it has verified the reported descriptions of tonal plateauing in Saramaccan. While this is not a new result, it lends valuable support to the transcribed tones for Saramaccan insofar as they are consistent with instrumental data. However, as we have seen, the instrumental data indicates that there are aspects of Saramaccan phrasal phonology which need to be further examined, specifically the possibility that particular words in phonological phrases are marked for primary and secondary accent and the way relative pitch levels are assigned to adjacent phonological phrases.

There are various ways in which the data collected for this study could be improved in future work. Perhaps the most important factor not controlled for in the elicited data was the segmental phonology of the words involved. It is well known that consonants can have effects on the F_0 of nearby vowels (see, e.g., Hombert (1978)). To the extent that segmental material was not controlled for, the figures presented in Table 8 may have not been the most accurate possible reflection of the real relative F_0 values of the different types of tones. Particular segments might also have affected some aspects of the pitch traces discussed in Section 8.[20]

Another area where improved data collection would be useful is that a larger number of phrases needs to be collected and analyzed in order to decrease reliance on impressionistic evaluation of pitch traces, as done here, and to, instead, come to conclusions about the pitch in phrases based on quantitative data.

References

Bruyn, A. (2002). The structure of the Surinamese creoles. In E. B. Carlin & J. Arends (Eds.), *Atlas of the Languages of Suriname* (pp. 153–182). Leiden: KITLV Press.

Daeleman, J. (1972). Kongo elements in Saramacca Tongo. *Journal of African Linguistics, 11*, 1–44.

Devonish, H. (1989). *Talking in Tones*. London: Karia.

Devonish, H. (2002). *Talking Rhythm Stressing Tone*. Kingston, Jamaica: Arawak.

Good, J. (2003). Tonal morphology in a creole: High-tone raising in Saramaccan serial verb constructions. In G. Booij & J. van Marle (Eds.), *Yearbook of Morphology 2002* (pp. 105–134). Dordrecht: Kluwer.

Good, J. (2004). Tone and accent in Saramaccan: Charting a deep split in the phonology of a language. *Lingua, 114*, 575–619.

Ham, W. (1999). Tone sandhi in Saramaccan: A case of substrate transfer? *Journal of Pidgin and Creole Linguistics, 14*, 45–92.

20. With respect to Figure 10, in particular, the surprisingly high F_0 of the first syllable of *kulé* 'run' could have been due to the generally observed fact that voiceless stops tend to raise the F_0 of a following vowel (Hombert 1978:79).

Hombert, J.-M. (1978). Consonant types, vowel quality, and tone. In V. A. Fromkin (Ed.), *Tone: A Linguistic Survey* (pp. 77–111). New York: Academic Press.

Hyman, L. M. (1978). Tone and/or accent. In D. J. Napoli (Ed.), *Elements of Tone, Stress and Intonation* (pp. 1–20). Washington: Georgetown University Press.

Kouwenberg, S. & Murray, E. (1994). *Papiamentu* [Languages of the World/Materials 83]. München: Lincom Europa.

Ladd, R. D. (1996). *Intonational Phonology*. Cambridge: CUP.

Liberman, M. I., Schultz, M., Hong, S., & Okeke, V. (1993). The phonetic interpretation of tone in Igbo. *Phonetica, 50*, 147–160.

Remijsen, B. (2001). *Word-Prosodic Systems of Raja Ampat Languages* [LOT Dissertation Series 49]. Utrecht: LOT.

Remijsen, B. (2002). Lexically contrastive stress accent and lexical tone in Ma'ya. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology VII* (pp. 585–614). Berlin: Mouton de Gruyter.

Rivera-Castillo, Y. (1998). Tone and stress in Papiamentu: The contribution of a constraint-based analysis to the problem of Creole genesis. *Journal of Pigin and Creole Languages, 13*, 1–38.

Rivera-Castillo, Y. & Pickering, L. (2004). Phonetic correlates of stress and tone in a mixed system. *Journal of Pidgin and Creole languages, 19*, 269–284.

Rountree, S. C. (1972). Saramaccan tone in relation to intonation and grammar. *Lingua, 29*, 308–325.

Voorhoeve, J. (1961). Le ton et la grammaire dans le Saramaccan. *Word, 17*, 146–163.

Tracing the origin of modality in the creoles of Suriname[1]

Bettina Migge
University College Dublin, Ireland

Creole TMA systems have figured prominently in debates on creole genesis. Bickerton argued that the similarities between creole TMA systems support the operation of a language bioprogram in creole formation. Lefebvre maintains that the similarities between the Haitian and Fongbe TMA system are evidence of the importance of substrate influence. While insightful, these studies are not conclusive. This paper therefore investigates the origin of two subsystems of modality in the creoles of Suriname. Based on a comparison of the strategies employed by three maroon creoles and their main substrate, and a consideration of early textual evidence, it argues that contact-induced and language-internal change contributed to the emergence and development of the modality system of the creoles of Suriname.

1. Introduction

Creole TMA systems have received a great deal of attention in the debate over the nature of creole formation ever since Bickerton (1981, 1984) argued that the significant

1. Earlier versions of this paper were presented at the workshop entitled "From Alada to Paramaribo 1651 to 1750: What happened to the language?" organized by the Dutch research group "A Trans-Atlantic Sprachbund? The structural relationships between the Gbe languages of West Africa and the Surinamese Creoles", April 2003, Wassanaar, Netherlands and at the annual meeting of the Society for Pidgin and Creole Languages (Boston 2004). The data come from fieldwork conducted in Benin and Suriname in 2002–2004 as part of the project "The influence of West African Languages on the tense/mood/aspect (TMA) systems of two Surinamese creoles" (Donald Winford, principal investigator). I would like to hereby gratefully acknowledge the funding of the National Science Foundation (NSF Grant #BCS-0113826). The insights presented in this paper are based on collaborative work between myself and Donald Winford. I would also like to thank the informants in Benin and Suriname/French Guiana for generously giving their time and H. Capo, D. Gagnon and J. Essegbey for helping with the collection of the Gbe data, and H. Capo, J. Essegbey and Enoch Aboh for insightful discussions about their interpretation. Thanks are also due to two anonymous reviewers for valuable comments and criticisms on an earlier version. All remaining errors are, of course, my own responsibility.

similarities between the TMA systems of (radical) creoles are evidence "for a linguistic-universals [bioprogram] explanation of creole genesis" (Singler 1990: viii). Two types of challenges have been leveled at this claim. First, research on the TMA system of individual creoles (see papers in Singler 1990; Lefebvre 1996; Winford 2000a/b) shows that the TMA categories and combinations or orderings posited by Bickerton's (prototypical) creole TMA system do not closely match those actually attested in creoles. Second, scholars have presented sound arguments in favor of the role of language-internal change (e.g. van den Berg 2001) and particularly substrate influence in the emergence of specific TMA-expressing elements in creoles. In relation to the latter, for instance, Corne (1983) argues that the semantics and use of completive *fin* in Isle de France Creole are a compromise between the semantics of French *finir* 'finish' and the completive category found in the Bantu languages that provided the major substrate input to the formation of Isle de France Creole. A similar case has been made for the sources of completive markers in Portuguese-lexicon creoles (Stolz 1987). Stolz argues that even the syntactic position of these markers in different creoles (pre-verbal or VP-final or both) can be attributed to the relevant substrates (e.g., Kwa languages in the Caribbean, Malay in the case of Papia Kristang). The most comprehensive argument for substrate sources of creole grammar, including TMA systems, can be found in the work of Lefebvre and her associates.

Lefebvre (1996, 1998) provides a detailed description of the TMA system of Haitian, its superstrate language French and one of its substrate languages, Fongbe, and an in-depth comparison of the Haitian system with that of French and Fongbe. With respect to modality, she argues that Haitian has three modality or irrealis markers, as she calls them. They are the definite future marker *ap*, the indefinite future marker *a va* and the subjunctive marker *pou*. The comparison reveals that the Haitian modality markers do not have direct counterparts in French. However, there are similarities in form and a few basic resemblances in meaning between the Haitian forms and elements in French periphrastic expressions. The Haitian element *pou*, for instance, is posited to derive from the French preposition *pour* occurring in periphrastic expressions such as (1) although French *être pour* and Haitian *pou* only share a tenuous semantic element in common, namely their irrealis meaning (Lefebvre 1998: 118–119).

(1) French *Jean est pour partir.*
'John is about to go' (Lefebvre 1998: 113)

The Haitian markers, however, match up very well in their semantics, syntax, and their combinatory possibilities with their Fongbe counterpart. Haitian *pou*, for instance, is argued to closely resemble the Gbe subjunctive marker *ní*. To Lefebvre, "these facts argue for the claim that the semantics of the TMA system of Haitian comes from its substratum languages" (Lefebvre 1996: 295–296) while the lexical forms seem to be derived from French lexical or periphrastic forms that share some semantic similarity with the Fongbe forms. This 'division of properties' suggests to Lefebvre that the process of relexification played a major role in creole genesis. Essentially, speakers of Fongbe copied the lexical entries of their (Fongbe) modality markers and rela-

beled the native lexical items with a phonetic string from French which comes from a somewhat semantically related element in French periphrastic constructions (Lefebvre 1996: 297).

While Lefebvre's work provides important insights into the nature of the Haitian TMA system and its relationship to its superstrate and substrate input, it is not entirely conclusive. First, it is just concerned with preverbal markers, i.e. grammaticalized strategies for expressing modal notions, and does not consider modal verbs and periphrastic constructions. This is quite surprising given that other creole functional elements (e.g. equative copula) have been found to originate from this kind of 'secondary construction' in the superstrate and/or the substrate (Migge 2002). Lefebvre's account explicitly acknowledges that French secondary constructions were involved in the emergence of the (lexical form of the) Haitian TMA markers but she does not discuss such constructions for Fongbe or Haitian. A possible second shortcoming to Lefebvre's account is that she does not consider diachronic data. These data are, however, important to provide a comprehensive account of the origin of creole TMA systems.[2]

The present paper attempts to shed light on the origin of creole TMA systems by investigating the emergence of two subsystems of modality in the creoles of Suriname. The investigation is based on a comparative linguistic analysis of modality in three maroon creoles and six Gbe varieties, and on a preliminary investigation of early historical documents (Goury 2003). The aim is to determine the role of the Gbe languages in the formation of these creoles and to show how input from both European and African sources, aided by universal principles of contact-induced change and language-internal change, shaped the grammar of these creoles. The paper suggests that many aspects of the creole modality system have their source in Gbe languages. At the same time, it is clear that they are in no way exact (or in some cases even close) replicas of the Gbe modality systems.

The comparative linguistic study follows the approach outlined by Thomason (1993: 287). According to this approach, a comprehensive analysis of language contact phenomena has to involve a careful analysis of the contact setting in which the contact occurred and an in-depth linguistic investigation of the affected subsystem of grammar in the languages that were involved in the contact setting and the resulting language.

The study focuses on the maroon creoles, Ndyuka (ND), Pamaka (PM) and Saamaka (SM) because they are quite conservative and unlike modern Sranan Tongo have undergone relatively little contact induced change.[3] The study focuses on varieties of Gbe (Aja, Waci, Gen, Xwela, Xwla, Maxi) because both sociohistorical evidence

2. To her defense, there may not be much historical data available yet.

3. A preliminary comparison of modality in the maroon creoles and Sranan Tongo (Winford 2000a) showed that the Sranan Tongo system differs in several important respects from that of the maroon creoles. These differences appear to be largely due to influence from Surinamese

(Arends 1995) and linguistic evidence (Migge 1998a/b, 2000, 2002, 2003; Smith 2001) suggest that speakers of Gbe played an important role in the formation of the plantation varieties from which all modern creoles of Suriname descend.[4]

The data for this study come from elicitations with selected native speakers and from recordings of natural conversations. The former data were elicited employing a modified version of Dahl's (1985) questionnaire. The questionnaire consisted of a number of sentences and short connected texts in French which were offered for translation to informants. Informants 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 habitual aspectual reference, a prompt sentence such as the one in (2) was used.

(2) [Q: What does your brother normally do after lunch? A:]
He WRITE letters.

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 or French when translating. In addition to providing equivalents of the English/French sample sentences, the informants were also encouraged to supply additional examples of their own, to evaluate differences in meaning between similar constructions, and to assess sample sentences constructed by the fieldworker/author. The elicited data were also evaluated and discussed in some detail with linguists who are native speakers of Gbe varieties.[5]

The conversational data were recorded by members of the different communities. These consultants were asked to make 60 minute recordings of mainly unguided and also one guided conversation with (conservative) members of their native village. The consultants who all had some training in linguistics or anthropology also transcribed and translated the recordings. The modality data from the recordings were also discussed with both the consultants and Gbe linguists.

Dutch on modern Sranan Tongo. More in-depth work on this issue is presently being conducted in the framework of the NSF grant.

4. The study focuses on several Gbe varieties because the creators of the predecessor(s) of the modern creoles of Suriname were speakers of different Gbe varieties. Moreover, such an approach provides insights into the nature of variation across the Gbe continuum. In the absence of a broad range of data from the very early period, a consideration of data from several modern creoles of Suriname was deemed necessary in order to get a better insight into the kinds of strategies that most likely existed at the time of creole formation.

5. An anonymous reviewer suggested that native speaker linguists are not a reliable source of information but did not explain why. I, however, found that discussions with native speaker linguists and the additional data supplied by them were very helpful in putting into perspective the collected data and the available literature on the Gbe varieties.

The framework used is modeled after typological studies of TMA systems such as Dahl (1985) and Bybee et al. (1994). The analysis focuses on semantic domains, e.g. necessity, and the various strategies, e.g. grammatical markers, modal verbs, adverbs etc., employed to express the meanings that are part of such a domain. With respect to each strategy used, the study aims to determine its dominant or prototypical uses and its secondary meanings, that is its contextually determined interpretations that arise from more peripheral uses. This is crucial because a close match between languages with respect to the nature of the strategies employed, their meanings and uses is a powerful argument for typological similarity between them. This would be strong evidence of continuity from the Gbe to the Surinamese varieties.

The study is divided into four parts. Part Two gives an overview of the strategies used to encode modal categories in the three maroon creoles and six Gbe varieties. Part Three discusses and compares the strategies employed to express potential mood and necessity in some detail and proposes a scenario for their emergence based on the comparative evidence and some historical data. The last part summarizes the findings and discusses their theoretical implications (on TMA marking see also Winford this volume).

2. Comparing modality in the creoles of Suriname and Gbe

Tables 1 and 2 provide an overview of modality categories in the maroon creoles of Suriname and Gbe.

A comparison of Tables 1 and 2 reveals several important similarities between the maroon creoles of Suriname and varieties of Gbe in the area of modality. They distinguish mostly the same kinds of semantic categories and express them with independent forms that precede the verb. They have closely similar meanings and distributions but different etymologies. The Surinamese forms derive from European languages such as English, Dutch and Portuguese. For reasons of space, the following discussion will focus on learned ability, potential and necessity.

3. Ability and possibility

The first striking similarity between the creoles of Suriname and Gbe is that in both language groups a mental or physical ability that requires special knowledge or learning is distinguished in the same way from other kinds of ability or possibility. Both employ a construction with the verb meaning 'to know', *sabi* in Ndyuka and Pamaka, *sá* in Saamaka, all derived from Portuguese *sabir*, and *nyã*, *nyɔ̃* etc. in Gbe. The examples in (3) illustrate:

(3)	PM	*A*	*pikin*	*de*,	*a*	*sa*(*bi*)	*suwen*.
		DET(sg)	child	DEM	s/he	know	swim

Table 1. Modality in three maroon creoles of Suriname

Forms PM[1]	ND	SM	Category	Meanings
		Learned Ability		
sabi	*sabi*	*sá*		Ability or skills acquired through learning or training
		Potential		
sa	*sa*	*sa*	Positive	**physical ability**
man	*poy*	*sa*	Negative	(Deontic) Ability subject to physical or natural law.
sa	*sa*	*sa*	Positive	**deontic (root) possibility**
man	*poy*	*sa*	Negative	Ability/possibility subject to moral or social law, involving situations under the agent's control
sa	*sa*	*sa*	Positive	**permission**
man	*poy*	*sa*	Negative	Deontic possibility imposed by authority (social, legal, etc.).
sa	*sa*	*sa*	Positive	**epistemic possibility**
man	*poy*	*sa*	Negative	Possible situations, or situations to the certainty of which the speaker is not committed.
kande	*kande*	*kande*		
		Necessity		
mu	*mu*	*musu*		**deontic necessity** or **obligation** "Existence of external, social conditions compelling an agent to complete the predicate action." Bybee et al. (1994: 177)
musu	*musu*	*musu*	(stronger)	
musu	*musu*	*musu*		
fu	*fu*	*u*	(strong)	
musu	*musu*	*musu*		**epistemic necessity** Inference based on sound evidence (prior knowledge, experience, etc.). Expresses a high degree of certainty on the speaker's part about some situation.
		Desire		
wani	*wani*	*kɛ*		Expresses speaker's desire and need.
		Need		
a(bi) fanoudu (fu)[2] *fanoudu*				Expresses speaker's need.

Note 1: PM=Pamaka, ND=Ndyuka, SM=Saamaka.
Note 2: Constructions are found in all varieties.

SM	*Di*	*mii*	*aki*,	*a*	*sá*	*wata*	*bunu.*
	DET(sg)	child	here	s/he	know	water	good
Xwla	*èvi*	*ɔ́*	*nyɔ́*	*tɔ́*	*lin*	*nyɔ́ɛ́nмé.*	
	child	DET	know	lake	swim	well	
Waci	*мəví*	*a*	*nyã*	*ə̀cì*	*fufu*	*yɔ̀nɛмə.*	
	child	DET	know	water	wash.RED	well	

'The child knows swimming (i.e. how to swim (well) (in the water/lake))'

Table 2. Modality in the six Gbe varieties

Forms			Category				Meanings
Aja	Gen	Waci	Xwela	Xwla	Maxi		
			Learned ability				
nya	*nya*	*nya*	*nyɔ*	*nyɔ*	*nyɔ*		
			Potential				
teŋu	*teŋ*	*tiŋ*	*hɛn teŋ(u)*	*tɛn*	*sixu*[6]	Pos.	**physical ability**
sɛnŋu	*teŋ*	*tiŋ*	*kpego su*	*kpeji*	*sixu*	Neg.	
teŋu	*teŋ*	*tiŋ*	*hɛn teŋ(u)*	*tɛn*	*sixu*	Posi.	**deontic (root) possibility**
sɛnŋu	*teŋ*	*tiŋ*	*kpego su*	*kpeji*	*sixu*	Neg.	
teŋu	*teŋ*	*tiŋ*	*teŋ(u) hɛn*	*tɛn*	*sixu*	Pos.	**permission**
sɛnŋu	*teŋ*	*tiŋ*	*kpego su*	*kpeji*	*sixu*	Neg.	
teŋu	*teŋ*	*tiŋ*	*teŋ*	*tɛn*	*sixu*		**epistemic possibility**
			Necessity				
ʍo a	*ʍo la*	*ʍo la*	*ʍo la*	*ʍo a*	*ʍo na*		**deontic necessity**
ʍo a	*ʍo la*	*ʍo la*	*ʍo la*	*ʍo a*	*ʍo na*		**epistemic necessity**
			Subjunctive				
ne	*ne*	*nə*	*ni*	*nɛ*	*ni*		**optative, hortative, jussive**
			Desire				
ji	*ji*	*ji*	*din jro*	*ka*	*jlo ba*		
			Need				
ʍo_ wudo	*hyã ji*	*hinyã*	*hyã din*	*kã*	*ba*		

In the maroon creoles, the verb 'to know' selects an activity verb or a noun that implies a certain activity, e.g. *wata*(*a*) implies the activity of swimming.[7] In Gbe, the verb *nyã*, *nyɔ̃* etc. also selects an activity verb in this construction. If it is an inherent complement verb (cf. Essegbey 1999) as in the case of 'swim', the verb is always accompanied by a noun (e.g. 'water', 'lake') because it is essentially the verb and the noun together that denote the activity. The order of the noun and the verb is reversed in Gbe and in some

6. Besides *sixu*, the conversational data from Maxi also contain the element *sikã*. It occurs much less frequently than *sixu* in the conversational data. The native informant never employed *sikã* in the elicitations.

7. There seems to be a small set of nouns that may be selected by 'to know' in this type of construction, e.g. *buku* 'book, read', *sitaati* 'street, how to act in town', *wagi* 'car, drive a car'.

varieties the verb is also reduplicated because the verb 'to know' in Gbe belongs to a class of verbs that trigger OV and OV.RED alternation (cf. Aboh to appear).

'To know' in the creoles and Gbe appears to be a lexical verb. First, it may also select other kinds of complements to express various functions. It may select a clausal complement headed by a complementizer/quotative (*taki* (EMC), *taa* (SM), *be* (Gen, Waci), *мə* (Maxi), *fan* (Xwela) etc.) to express the notion of 'to know that …' or an NP complement to express 'to know something'.

Second, verbhood tests for these languages (cf. Lefebvre & Brousseau 2002: 101–107; Aboh 2004: 168–170) provide positive results. Like other lexical verbs in these languages, 'to know' can, for instance, be predicate-clefted in this function (4).

(4) PM *Sabi a sabi wata/suwen so?*
know she know water/swim so
'She KNOWS how to swim that well?'
Waci *Nyã мəvi a ye nyã əci fufu ləkə?*
know child DET PRE know water swim.RED like.that
'Does the child KNOW how to swim like that?'[8]

Moreover, it may be combined with the imperfective marker in the creoles and appear in the progressive construction in Gbe, like other lexical verbs in these languages (5).

(5) PM *A e sabi wata/swen.*
she IMPF know water/swim
'He is getting to know how to swim'
Waci *мəvi a le eci fufu nyã ŋ.*
child DET(sg) COP water swim.RED know PART
'The child is knowing, i.e. getting to know, how to swim'

Finally, the VP of 'to know' may be elided in the maroon creoles. In the case of the Gbe languages, the VP complement may not be deleted, most likely because *nyã* etc. are inherent complement verbs (cf. Aboh to appear; Essegbey 1999) which always requires the expression of the NP. The NP may, however, be extraposed.

(6) PM A: *Da a sabi swen/wata?* B: *Ya baa, a sabi.*
then she know swim/water yes POL she know
A: 'Does she know how to swim? B: Yes, she knows'
Waci A: *мəvi a nyã əci fufu a?*
child DET know water swim.RED QP
B: *En e nyã *(eci fufu).*
yes she know water swim.RED
'A: Does the child know how to swim?
B: Yes, she knows how to swim'

8. Capitalization of entire words indicates special emphasis.

Waci *En, eci fufu ye e nyã.*
yes water swim.RED FOC she know
'Yes, she knows how to SWIM'

Most of the Gbe varieties and the EMC also use a single preverbal form to express the semantic notion of physical ability, deontic (root) possibility, and permission. In the maroon creoles, the form *sa* is used to convey all these senses, see examples (7)–(9).

(Physical Ability)

(7) ND *A taanga, a sa diki wan ondoo kilo.*
he strong he can lift one hundred kilo
'He is strong, he can lift one hundred kilos'

(Permission)

(8) PM *Mi mama no wani fu en pikin, en umanpikin, go libi anga den sama dati, ma en manpikin sa go libi.* (PM1)[9]
my mother NEG want for her child her daughter go live with DET(pl) person that but her son can go live
'My mother doesn't want for her daughter to live with those people but her sons may live with them'

(Root possibility)

(9) PM *Den sa kon puu u ma den ná o puu u.*
they can come remove us but they NEG FUT remove us.
'They [Surinamese government] may come to remove us [from the gold- mining area] but they won't (be able to) remove us' (PM1)

Most of the Gbe varieties (Gen, Waci, Maxi, Xwla) also employ a single form to convey all these meanings, as shown by the examples (10)–(12) from Xwla, but the forms are not the same in all varieties (Table 2).

(Physical Ability)

(10) Xwla *é jɔ́ àsú. é tɛ́n kɔ́ kilo kɔnwèwì.*
he be man he can lift kilo 100
'He is very strong, i.e. he acts like a man should. He can lift 100 kilos'

9. In the case of examples taken from the conversational data, I provide the name of the variety and the tape number: in round brackets following the example, e.g. (PM 1).

(Permission)

(11) Xwla *èvi ɔ̀ tɛ́n nɔ̀ fí xwèsà xèɔ̀.*
child DET can stay here night this
'The boy may stay here tonight'

(Root possibility)

(12) Xwla *Vɔ̀ é tɛ́n lutter ná mɛ̀-bú yá é tɛ́n ún*
but he can fight PREP person-other EMPH he can NEG
kpé dó ɛ̀mɛ̀-bú wú.
associate PREP person-other skin
[Talking about the presidential campaign:] 'But he can fight for someone else, he can get together with someone else' (Xwla 4)

Note that in several Gbe varieties, such as Aja and Vhe varieties (e.g. Waci, Anlo), and to a lesser extent in Gen, this form is always combined with the future potential marker (13).[10]

Essegbey (p.c. November 2003) for Anlo and Capo (p.c. March 2004) for Waci argue that it is this marker that contributes the possibility interpretation.[11] *Tɛn* in Xwla, *teŋu* in Xwela and *sixu* in Maxi are generally not combined with (*l/n*)*a* in non-future contexts (13) since the future marker does not appear to have a modal meaning in these varieties (Aboh 2004: 158–164 but see also Lefebvre and Brousseau 2002: 91ff. for a different point of view).

(13) Waci *Yə́bə́ àʍù-wó sə́ŋú ŋ̀tɔ́; yə́ á tíŋ́ á kpà àgɔ́nə́*
his tooth-PL strong very they FUT can FUT cut coconut
kú-wó jró. (Capo p.c. 2003)
with-PL will
'His teeth are so strong that he can peel a coconut with them'

10. The term Vhe comes from Capo (1988). The cluster is more widely known by the name of Ewe.

11. Evidence suggests that (*l*)*a* in these varieties conveys a sense of (future) intention. To convey a sense of certain future, (*l*)*a* has to be combined with a periphrastic construction that expresses certainty (i).

(i) Waci *Mə̃ kà ʍə́ ə-jí bə́ Jáã̄ á yì Flã̄sə́-dù mə́ lə́ kɔ̃siʍá*
I bet LOC it-on that Jean POT go France-village in LOC week
nə́ gbɔ́nɔ mə
REL come in
'I am sure that Jean will definitely go to France next week.'

(14) Maxi *Nyavi ɔ, (e) sixu nɔ fi ogbadannu.*
boy TOP he can stay here evening
'The boy can (i.e. is authorized to) stay here'

In Xwela, the situation is a little bit more complex. In positive constructions, there is variation between the element *teŋ(u)* 'can' (15a) and a bi-clausal construction involving the verb *hɛn* 'carry' (15b) to express physical ability, possibility, permission and epistemic possibility. The Xwela consultant does not see any clear difference in meaning between these two options. In other varieties of Gbe, such as Gun and Waci, the construction with *hɛn* implies a lesser commitment to the truth and lesser certainty than the one involving *teŋ(u)* (Aboh p.c. 2004; Capo p.c. 2003).

(15) a. Xwela *vi lɔ teŋ wa jɛ ɯe hwedo.*
child DET can be fall LOC night

b. Xwela *vi lɔ hɛn ɔ, e la wa ɯe hwedo.*
child DET carry PART he FUT be LOC night
'The child can remain here tonight'

4. Negative possibility

Another interesting similarity is found in the expression of possibility in negative contexts. In such constructions, Pamaka and Ndyuka and several of the Gbe varieties (Aja, Xwela, Xwla) employ forms that are different from those used in positive contexts (16)–(18). Pamaka (and Aluku) employs the element *man*, Ndyuka *poy*, Aja *sɛnŋu* and Xwela and Xwla use *kpego/wu/ji*. The conversational data from Xwela show that *su* (<*isu* 'prohibition') can also be employed to express negative possibility and negative permission.

(Negative physical ability)

(16) PM *K., da i no man leysi a beybel moo?*
name, then you NEG can read DET(sg) bible anymore
'K, then are you no longer able to read the bible, i.e., hold mass'
(refers to the fact that K.'s vision has considerably deteriorated)
(PM 6)

ND *Ape yu á poy pasa anga a lay a boto.*
there you NEG can pass with DET(sg) load at boat
'There you cannot pass with a loaded boat' (ND 1)

Xwla *sán míè nɔ́ sɔ́n klàkú sán klàkú é hàn-làn*
before we HAB go Klakou before Klakou PART pig-meat
flan lókpó-nù ɔ̀ ún kpéwú á ɯù flan-lókpó. (Xwla 3)
franc one-for you NEG can[12] FUT eat franc-one

'When we used to go to Klaku before, it was pork for one franc, you could not eat for one franc'

(Negative permission)

(17) PM *A boy án man tan ya tide neti.*
DET(sg) boy NEG can stay here today night
ND *A pikin á poy tan ya tide neti.*
DET(sg) child NEG can stay here today night
'The child cannot stay here tonight'
Xwla *èvi ɔ̀ ún kpéjí á nɔ̀ fí hwèsà xèɔ̀ ò.*
child DET NEG can FUT stay here night this NEG
'The child cannot (i.e. absolutely not) stay here tonight'

(Negative root possibility)

(18) ND *Mi á poy de na a kiiman se.* (ND 2)
I NEG can COP LOC DET(sg) killer side
'[why he won't vote for the party:] I cannot be/work together with a killer'
PM *So wan libi so wan libi ná o man libi.* (PM 6)
so a life so a life NEG FUT can live
'[about troubles with wife:] Such a life, such a life cannot be lived'
Xwla *mɛbú è cán ká kplàn tò ɔ̀ àzan-mè*
person-other PL also want accompany country DET night-in
kènná ɔ̀yɔ̀ ègá ɔ ún gɔ kpéwú ká kèbɛ́
when you chief DET NEG again can want open-mouth
gànxɛn-mɛ lè àzan-mɛ ò-à.
moment-in LOC night-in NEG
'Other people also want to govern the country in the night, you, the boss, you cannot say anything now in the night' (Xwla 4)
Xwela *yɛlɛ nɔ do na mi yɛ tete lɛ mi hɛn lɔ mi su*
they IMPF accuse PREP us COR even PL we want PART we ??
gbɛ ɛ?
refuse QP?
'It is them who are accusing us, but if we want we are able to refuse, right?' (Xwela 2)

Saamaka patterns with several other Gbe varieties (Gen, Gun, Waci, Maxi) that employ the same modal in both positive and negative contexts (19)–(21).

---

12. The consultants who transcribed the conversational data consistently glossed *kpego/wu/ji* with French 'pouvoir', i.e. English 'can'. An anonymous reviewer and Aboh (p.c. 2004) suggest that this gloss may not be entirely adequate. For want of a better gloss, I stick to 'can' for now, but see also the discussion on its structure, categorial status and meaning below.

(Negative physical ability)

(19) SM *Mé sa ko tide moo, mi fii siki tide.*
I-NEG can come today more I feel sick today
'I cannot come tonight/today because I am/feel sick'

Gen *Docteur ke gba va ya e ya gba nyran wu.*
doctor who again come DEM he DEM again be bad body
'The doctor who arrived, as for him, he is even worse'
o ma teŋ mi o.
you NEG-FUT can swallow NEG
'You cannot even swallow it (the medicine)' (Gen 3).

(Negative permission)

(20) SM *Á sa fika ku mi di ndeti aki.*
NEG can remain with me DET(sg) night here
'The child/he cannot stay here tonight'

Waci *ə̀vié àmə́-bú mə̀ vi nyi yé vìé ma*
child person-other POSS child COP COR child-PART NEG-FUT
t[iŋ] a nɔ gbɔ.
can FUT stay at
'The child, it's a child of someone else and the child, can't he stay with her?' (Waci 2)

(Negative root possibility)

(21) SM *Mé a(bi) moni, nou mi á sa go a booko di*
I-NEG have money now I NEG can go LOC break DET(sg)
dia.
day
'I don't have money, thus I cannot go to the party'

Waci *Mi yí T. A. gbó T. A. be yé ma teŋ yí*
we go name name at name name say she NEG-FUT can go
a wa. yé мəká tutú kpo lə́ xɔ mə́ ku víe ó.
FUT come she one IEDO only COP house in with child PL
'We went to T. A.'s and T. A. said that she cannot go. It's her alone in the house with the children' (Waci 1)

In Waci, Gen, Maxi (and Fon) *kpe-ji/go* and *kpe-wu/ŋu* may also be used to express negative ability (22).

(22) Waci A: *Kōfí á tíń á kɔ́ àgbà mā à?*
name FUT can FUT lift load DEM QP
'Can Kofi lift that load?'

B: *Mə́ á kpə́-ə́jí ò, ʍóá mə́ ʍù nú ə́gbɛ̀ ò.*
NEG FUT can NEG because NEG eat thing today NEG
'He cannot lift it because he did not eat anything today' (Capo, p.c. 2004)

In all varieties of Gbe *kpe-ji/go* and *kpe-wu/ŋu*, just like *man* and *poy* in Pamaka and Ndyuka, are also used in positive declarative constructions and in questions to convey emphasis (23B, 25B) or to express a challenge (23A, 24, 25A).

(23) Xwela A: *ɔ̀yɔ̀ fé aá kpégò lō sɔ̄ ɔ̀?*
you also FUT can for go QP
'Can you also go?'
B: *lāá kpégò dǐ!*
FUT can very
'Of course, I can!' (Capo, p.c. 2003)

(24) Xwla *fínɔ̀ wé à kpé-jí á gɔ̀ bá à kpé ɔ̀ŋkúsì do*
where FOC FUT can FUT still come FUT can eye put
suklu wú. (Xwla 2)
school on
'[talking to a girl who had a child while still at school] Where can you still go, how will you be able to take care of your studies?'

Maxi *Emi ma kpewu o na xwenuʍɔ ǎ ani utu emi nɔ*
you NEG can and FUT speak NEG why cause you HAB
ji fi sukpɔ ɛ? (Maxi 2)
give.birth child many QP
'[If] you cannot talk, why do you have that many children?'

(25) PM A: *I man sikiifi a biifi gi mi?*
you can write DET(sg) letter for me
'Are you (really) able to write this letter for me?'
B: *Iya, mi man sikiifi en gi i oo!*
Yes I can write it for you EMPH
'Yes, I am able to write it for you!'

5. Epistemic possibility

In the maroon Creoles, the category of epistemic possibility, that is, possibility based on the speaker's subjective evaluation of the situation is either expressed by *sa* or by the adverb *kande* in conjunction with the future marker *o* or with *sa*, see examples in (26).

(26) PM *Goon yuu, a pasa tuu, a tuu, ma a sa koti ete.*
field hour it pass true it true but it may cut yet
'The time for preparing the field has truly passed, but it may still be cut' (PM 2)

PM *J. án sa go a faansi taa wiki.*
J. NEG may go LOC France other week
'J. may not go to France next week'

ND *Kande alen o/sa kai tide bakadina.*
maybe rain FUT/may fall today evening
'It may rain this evening'

PM *Den ná e kon, neen mi puu en, kande sama sa go de.*
they NEG IMPF come then I remove it maybe person may go there
'They [animals] were not coming, then I removed it [gun], someone may go there' (PM-NSF)

It is also possible to employ the construction *a sa de taki* to convey epistemic possibility (27).

(27) PM *A sa de taki a o go a Faansi dyonson.*
it can COP say she FUT go LOC France soon
'It's possible that she'll go to France soon'

In the Gbe varieties, epistemic modality is generally conveyed by the element *téŋ* etc. which is also used to express ability, permission and root possibility (28). As in the creoles, epistemic modality can also be conveyed by a periphrastic type construction as in (29) though.

(28) Gen *Jan lá téŋ yi yovódé wetrí ʍe a mɛ.*
Jean FUT can go Europe month other DET in
'Jean may go to Europe next week'

Gen *Jan (lá) téŋ nɔ axómɛ fífijén faa.*
Jean FUT can COP house now easily
'Jean may be at home now'

(29) Waci *ə́ tíŋ́ nyí bə́ Kōfí ʍù-nà nú cú sã́.*
it can COP that name eat-HAB thing all before
'It is possible that Kofi was greedy before' (Capo, p.c. Nov. 2003)

Waci *Ta mə nyɔ́ o a, Kɔ̀jó jó vɔ̀.*
head NEG be.good NEG PART name go finish
'Maybe Kojo is already gone (and I wish he is not gone)' (Capo, p.c. March 2004)

6. The categorial status of the modality elements

The categorical status of the elements employed to express physical ability, deontic and epistemic possibility and permission in Gbe and the maroon creoles is somewhat difficult to pin down in most cases. Some elements, notably *poy* in Ndyuka, *man* in

Pamaka (and Aluku), *sixu* in Maxi, *teŋ*(*u*) in Gen and *tiŋ* in Waci appear to be modal verbs rather than auxiliaries because their VP complement may be deleted:

(30) PM A: *Luku a gaan beki, (n)ku a gaan*
look DET(sg) big container look DET(sg) big
beki de.
container there
B: *A ná o man.*
he NEG FUT able
A: *A o man.* (PM 12)
he FUT able.
'A: Bring me that big container.
B: He cannot (carry it).
A: He can (carry it).'

ND A: *A be o, sa wani booko den sowtu sani de.*
he PAST FUT may want break DET(pl) sort thing there
Ma a poy? A á poy!
but he can he NEG can
B: *Mmm! A á poy!* (ND 1)
EXCL he NEG can
'A: He wanted to destroy all these things (the power of maroons). But could he? He couldn't!
B: He couldn't'!'

(31) Maxi *O na lɛ ʍɔ emi na do alũmɛ nu,*
you FUT again say you(EMPH) FUT plant dry-season thing
ehi ũkpŏgĕ ɔ, o sixu a! (Maxi 2)
you(EMPH) alone PART you can NEG
'You'll say again that YOU'll plant for the dry season, YOU alone, you cannot!'

Waci *Gàmè nə mə nyì ʍəkajə́ a mə tiŋ fùʍu kaba vɔ*
when that I COP young.man DET I can run fast but
fifia nyə má tiŋ ò.
now I NEG-FUT can NEG
'When I was young I could run very fast but now I cannot (anymore)'

However, they do not seem to be fully lexical verbs either because they cannot be predicate-clefted, see for instance example (32) from Pamaka.

(32) PM **Na man a án man tya en.*
PRE can he NEG can carry it
'He CANnot carry it'

Moreover, they cannot be modified by the imperfective marker or any other tense or aspect marker – these markers only modify the main verb. However, they may combine

with tense and aspect markers to form complex temporal notions that modify the main verb of the construction (33).

(33) PM *Da i mu man taagi a kode-nəmə.* (PM 12)
then you must can tell DET(sg) telephone-number
'Then you should be able to tell (them) the telephone number'

ND *Di a pasi tapu ya, a den sisa ná o poy*
when DET(sg) path close here PRE DET(pl) sister NEG FUT can
waka go anda. (ND 4)
walk go over.there
'When the path closes here, the women won't be able to go there anymore'

Maxi *Ani emi ka nɔ blo ehɔ̃mɛ dĩ bo na sixu hɛ̃ egbɛ*
what you ? HAB make today now COR FUT can hold life
emitɔ̃ ʍo ete na?
your COP upright QP
'What do you usually do to make a living?' (Maxi 1)

Gen *e la ke teŋu wɔ dɔ a.*
he FUT CF can make work DET(sg)
'He would have been able to do the work.'

Based on the available data, it is not clear whether *teŋ*(*u*) in Aja and Xwela, *tɛn* in Xwla and *sɛnŋu* in Aja function entirely in the same way as the above elements or whether they have more verbal or marker qualities.

Sa in the maroon creoles behaves much more like an auxiliary than a verb. It cannot be predicate-clefted, the VP following it cannot be deleted and it can only combine with very few T/A markers such as the relative past time marker *be* to express hypothetical possibility or ability (34).

(34) PM **Na sa i sa pee poku tide.*
PRE can you can play music today
'You CAN play music today.'

PM A: *Mi sa pee poku tide? B:*Iya, i sa.*
I can play music today Yes, you can
'A: Can I play music today? B: Yes, I can.'

ND *Te yu à man daay i konde moo, da den be*
when you NEG can turn your country more then they PAST
sa yeepi i baka. (ND 4)
may help you back
'When you cannot handle your country any more, then they would help you out/again.'

It also combines with the imperfective marker *e*. Note, however, that *e* follows *sa* but precedes lexical and modal verbs. This may be taken as a further piece of evidence that *sa* is not a lexical or modal verb but rather auxiliary.

(35) PM *Efu a M. be de a opu ya, i be sa*
If PRE name PAST COP LOC up.river here you PAST may
e wooko. (PM 12)
IMPF work
'If it's M. who was living in the up river [village] here, you could be working.'

The status of *kpeji/go/ŋu/wu* is not entirely clear. Structurally, it does not seem to be a word but rather a phrase that has the structure in (36).

(36) V[*kpe 'reach'*]—PRO[*e 'it'*]/NP—Postposition[e.g. *ji 'top, summit', ewu, eŋu 'body'*]

In addition, tense and aspect markers typically follow it (37), i.e. they directly precede the main verb, or are marked on *kpe* and the main verb (Waci 19).

(37) Xwla *ún má lɔ̀ bé ò yé ún kpéwù á bá.*
he FUT-NEG say that no he NEG can FUT come
'He will not say that he cannot come.' (Xwla 4)
Xwla *ún kpéjí lé ká zùn yé ò.*
NEG can IMPF want hurt them NEG
'They cannot be wanting to hurt them' (Xwla 3)

It seems best to classify it as a phrasal construction that is akin to the 'make it' (e.g. Will they make it to the meeting?) construction in English (Aboh p.c. June 2004).

7. The emergence of possibility and ability in the maroon creoles of Suriname

The comparison of the means of expressing ability and possibility in the maroon creoles of Suriname and varieties of Gbe revealed close similarities between the two language groups suggesting that the Gbe languages had an important impact on this area of grammar in the creoles.[13] The similarities are particularly striking with respect to the expression of ability acquired through special teaching or learning. The most likely explanation would be that speakers of Gbe established an interlingual identity between the Portuguese derived word *sabi* and their native Gbe item *nyã* etc. in contexts in which *sabi* was used as a main verb selecting an NP complement to express the notion of 'to know something' (see (38) below). As a result of this association, the semantic and syntactic properties of Gbe *nyã* etc. were projected onto *sabi*. *Sabi* thereby came to select not only NP but also sentential and VP complements and to express additional notions such as 'to know that …' and 'to know an activity …' besides 'to

13. An anonymous reviewer suggested that an important evidence in favor of L1 influence would be that the Gbe and creole elements match up in terms of their semantic and syntactic properties (e.g. categorial status). Note that while such a perfect correspondence would be nice, it is extremely rare, if not impossible to find (cf. Thomason & Kaufman 1988).

know something' in the creoles (cf. 2). Given that Portuguese *sabir* expresses similar meanings, it seems likely that both substrate and superstrate influence conspired to give rise to the semantic and syntactic properties of *sabi*/*sá* in the creoles.

(38) Gen *Nyɛ́ nyɛ̃ mu nyã do na.*
I (EMPH) I (EMPH) NEG know depth DEM
'Me, I don't know the depth of that problem' (Gen 1)

PM *Mi sabi a sama.*
I know DET(sg) person
'I know the person'

The matter is somewhat less clear in the case of the other elements. The form *sa*, usually represented as *za* or *zal* in the early documents (SN), most likely derives from the first and/or third person singular present form of the Dutch modal auxiliary *zullen*. In Dutch, *zullen* expresses an obligation or a prohibition (39), or a probability when it occurs together with the adverb *wel* 'surely, certainly' (40) (ANS).[14]

(39) Dutch *Je wou het hebben, nou zal je het ook opeten.*
you want it have now should you it also finish.eat
'You wanted to have it, so now you also have to eat it all' (ANS – 53)

(40) Dutch *Pet zal wel slagen.*
name should surely succeed
'Pet will probably succeed' (ANS – 43a)

In the texts written in the early plantation creole (SN), *sa*, however, conveys future time reference. In combination with adverbs expressing possibility or probability, *sa* conveys uncertain future and epistemic possibility (41) and without such markers it expresses a more definite future time reference (42).

(41) SN *Zomtem a za kom jusse na.*
Maybe he ?? come just now
'He may be back any moment'
(Arends & Perl 1995: 146, van Dyk 1765 cited in Goury 2003)

(42) SN *Mi za kry da tem joe go wee.*
I FUT cry that time you go away
'I'll cry when you leave'
(Arends & Perl 1995; van Dyk 1765 cited in Goury 2003)

Unlike its Gbe counterparts, *sa* never functions as a marker of physical ability, permission or root possibility in the early texts. The few instances of root possibility, permission and ability are expressed by the modal verb *kan* (43)–(44) (Goury 2003) which is also still widely used in this function in modern Sranan Tongo (cf. Winford 2000a).

14. Note that it is relatively difficult to provide a handy English equivalent for Dutch *wel*.

(43) SN *Den no kann holli lieki den tarrewan.*
they NEG can hold like DET(pl) others
'[speaking about pregnant women:] They cannot work like the others'
(Arends & Perl 1995; van Dyk 1765 cited in Goury 2003)

(44) SN *Da misi takki offe mastra plessi a kan go lange da boote.*
that woman say if master please he can go with that boat
'The woman said that if the master wants to, he can go with the boat'
(Arends & Perl 1995; van Dyk 1765 cited in Goury 2003)

The currently used definite future marker *o* (< English *go*) (45) does not appear in the texts before the end of the 18th century (Goury 2003).

(45) SN *Effi koure kissi datti, a go pori.*
if cold get that it FUT spoil
'If it gets cold, it'll spoil' (Schuman 1783 cited in Goury 2003)

Given these early data, one possible scenario would be that *sa* initially emerged as a future-marking element. The most likely scenario would be that native speakers of Gbe established an interlingual identity between *sa* and the native Gbe future marker. As a result, *sa* most likely came to encode notions ranging from a more definite future to an uncertain future. *Sa* then became restricted to expressing potential and uncertain future and epistemic possibility, and (towards the end of the 18th century) definite future came to be encoded by (*g*)*o*, which van den Berg (ms.) argues is derived from a construction involving the imperfective marker *de* and the motion verb *go*. In Sranan Tongo, *sa* still today largely expresses only the former meanings (cf. Winford 2000a/b). However, in the maroon creoles it came to also convey deontic possibility, permission and physical ability, that is, meanings that in the early texts are expressed by *kan* (43)–(44). Based on the available evidence, it seems that *sa*'s development into a potential marker in the maroon creoles must have come about due to several processes. First, substrate influence is likely to have played a role. Given that in Gbe epistemic possibility is expressed by the same markers as deontic possibility, permission and physical ability, *sa*'s extension to the latter contexts may have been due to influence from Gbe. Second, it is also likely that *sa*'s extension was reinforced by processes of language-internal change. Third, at least *sa*'s extension to convey permission could have also been due to influence from Dutch *zullen*. At this point, it is not entirely clear what the exact role of these sources were and whether *sa*'s extension to modal contexts occurred only after the emergence of *o* as a definite future marker or whether *sa* already expressed modal meanings (in the maroon creoles) prior to the end of the 18th century. Research on this issue is currently being carried out (cf. Goury & Migge forthcoming a and b).

While there seems to be a nice match between several Gbe varieties (Xwela, Xwla, Aja) and the maroon creoles Ndyuka and Pamaka in that they all use distinct markers to express negative and positive possibility, ability and permission, it is not clear that this match is entirely due to substrate influence. According to early textual evidence examined by Van den Berg (2001), the auxiliary *man* seems to have emerged as a negative physical ability marker due to a process of reanalysis from the noun *man*. Once reanalyzed, it spread to the other contexts of possibility due to a process of extension and replaced *sa* in negative contexts in Pamaka (and Aluku). However, it also seems somewhat unlikely that this would have been entirely an internal development and that the similarities between Gbe and Pamaka and Ndyuka are purely accidental. First, the distribution and function of *man* and the negative deontic possibility, permission and physical ability items in Gbe are very close. Second, there is also a semantic similarity between the original forms. *Man* was reanalyzed from a construction expressing lack of physical strength literally meaning 'he is not a man for X' (46). The items used in negative contexts in Gbe also derive from phrases conveying physical strength (47). This suggests that the Gbe patterns probably spawned the reanalysis of *man*.[15]

(46) SN *Mi no man va hoppo datti.*
I NEG man for lift that
'I am not strong enough to lift that (lit.: I am not man for lift that).'
'I cannot lift that' (Schumann 1783: 185 cited in van den Berg 2001: 249)

(47) Aja *sɛ̃-ɛ̀ŋu*
be hard/be strong-body
'be strong'
kpé-é-(e)ŋú-(ci)
reach-its-body-tree/body
'be as strong as'
Xwla/ Waci *kpé-é-ji*
reach-its-summit/top
Xwela *kpé-e-gò*
reach-its-body
Maxi *kpé-e-ewú*
reach-its-body
'be strong' (Capo, p.c. 2003/Aboh p.c. 2004)

15. One reviewer suggested that the creole construction may have also been spawned by the Gbe 'daring construction' which according to Aboh (p.c. 2004) is *Kofi jɔ gbeto*. 'Kofi acts (like) a human.' (Gungbe). However, based on the available data, it is not clear whether this construction is widely used in Gbe to express notions such as ability, possibility, etc. More research is necessary on this construction.

The case of *poy* in Ndyuka is somewhat less clear. According to Smith (1987:222) it derives from the third person singular form *pode* of the Portuguese verb *poder* meaning 'can'.[16] However, *poy* in Ndyuka does not seem to be entirely modelled on the Portuguese element since its distribution does not match up with *pode* in Portuguese. Unlike *pode* but like *man* in Pamaka, *poy* in Ndyuka is employed with all persons and mainly used in negative contexts to express physical ability, deontic possibility and permission. In positive contexts it expresses emphasis and conveys a challenge. The close match in distribution and function between *man* and *poy* suggests that *poy*, like *man*, must have been largely modelled on the Gbe elements that express negative physical ability, possibility and permission (47). It seems that *poy* co-existed with *man* in the early maroon and plantation varieties.[17] Today it is regularly used as a negative potenial mood marker by members of the Ndyuka community. Alukus and Pamakas do not readily use it because for them it is largely associated with a Ndyuka ethnic identity.

8. Necessity

Another important similarity between the creoles of Suriname and the varieties of Gbe is that they both use the same form to express positive and negative strong and weak obligation, as illustrated in examples (48)–(49). Strong obligation refers to the "existence of external, social conditions compelling the agent to complete the predicate action" (Bybee et al. 1994:177) while weak obligation conveys the speaker's recommendations or advice. In Ndyuka and Pamaka these two senses of obligation are expressed by the form *mu*, Saamaka employs *musu* (< English *must*), and in the Gbe varieties, strong and weak obligation are conveyed by *мо* (*l/na*) (*мо* < *мо* 'have, possess', *l/na* < *l/na* 'future').[18]

(48) PM *Mi anga yu, a moyti u mu meki a den baka.*
me with you it effort we must make LOC their back
'[talking about relationship to Europeans:] Me and you, we must make an effort to keep up with them' (PM 1)

16. Smith argues that the path of development was as follows Portuguese *pode* > *poli* > Ndyuka *poy*.

17. Schumann's Saamaka dictionary from 1778 lists *poli* (> *poy*) and provides the following sample sentence: *mi no poli va go na matu tide* 'Today I am not able to go to the forest.' (Schumann 1778:97)

18. One anonymous reviewer suggested that *мо* (*l/na*) was only employed in Western Gbe varieties and that Eastern Gbe (e.g. Fongbe varieties) used another form but did not suggest a form. My research on Maxigbe and discussions with native speakers of other Eastern Gbe varieties etc. suggests that *мо* (*l/na*) is also regularly used as an obligation marker in these languages.

SM *Yee i kɛ di moni fi i, nou tide ndeti i*
if you want DET(sg) money for you then today night you
musu ko a mi.
must come LOC me
'If you want your money, you must come to me tonight'

Maxi *Egbɛtɔ ʍo na nu esĩ.* (Maxi 2)
human must FUT drink water
'Humans must drink water'

(49) PM *I án mu membe taki ná wan sani di a e du*
you NEG must believe that NEG one thing REL he IMPF do
fu a go anga baka [...]. *A soso a fesi a e*
for he go with back FOC only DET(sg) face he IMPF
gwe. (PM 1)
go
'You should not believe that he [European] is doing anything that makes him go backwards, he only moves forwards'

SM *Di sembɛ dɛ a musu suku muyɛɛ bifo a gaandi.*
DET person there he must search woman before he old
'That man should find a wife before he is old'

Maxi *O ʍo na kple akwɛ tuwe.*
you must FUT collect money your
'You should save your money'

However, if it is necessary to distinguish between the two, a periphrastic construction may be used to augment the degree of obligation in Gbe (50).

(50) Waci *ə́ hyā bə́ ʍəvi ā ʍó lá kplɔ̀ xɔ a mə̀.*
it need that child DET must FUT sweep room DET in
'It is necessary that the child must clean the house' (Capo p.c. 2003)

In the creoles, strong(er) obligation may be conveyed by *a*(*bi*) (*f*)*u* (*habi* < English *have*, *fu* < English *for*') in all varieties and by *musu fu* in Ndyuka and Pamaka, and by *musu u* in Saamaka. Compare the examples in (51).

(51) PM *Da ala fasi, a abi fu kon a ini a famii ini.*
then all fashion he have for come LOC in DET(sg) family in
'[talking about a former avenging spirit:] In any case, he has to come (back) among the family' (PM 1)

PM *Di mi be nyoni, mi be musu fu kiin a osu.*
when I PAST small I PAST must for clean DET(sg) house
'When I was small I had to clean the house (and did it)'

In negative sentences, *mu*(*su*) and *ʍo* (*l*/*na*) express admonition or forbidding constrained by moral law (52).

(52) PM *Den lanti án mu koli den foluku fu den.*
DET government NEG must fool DET(pl) people POSS them
'The government must not fool its people.' (Sanna, p.c.)

Waci *PC mə́ ʍó lá flù àgblədətɔ́ wó ò.*
communiste.party NEG must FUT fool farmer PL EMPH
'The communist party must not fool the peasants' (Capo, p.c. 2003)

To express an unfulfilled past obligation, the creoles combine *mu* and *musu* with the past time marker (53).

(53) PM *Mi be mu baka wan kuku tide, ma ten án be de moo.*
I PAST must bake a cake today but time NEG PAST COP more
'I was supposed to bake a cake today but there wasn't any time left'

SM *Mi bi musu yasa wan kuku tide ma mé bi a tin u yasa en moo.*
I PAST must bake a cake today but I-NEG PAST have time to bake it more
'I was supposed to bake a cake today but I did not have time left to bake it'

In some Gbe varieties, such as Gen and Aja, *ʍo* (*l/na*) is combined with the irrealis or hypothetical-marking element *ke* (54) while in other varieties such as Xwla, Xwela and Maxi the future marker is combined with the perfect marker, *nɔ* in Xwla, (55) to express an unfulfilled obligation.

(54) Gen *Mu ʍó lá kẽ tɔ botokyɛn égbe, vɔa nyɛ mú vo o.*
I must FUT IR bake donuts today but I NEG free NEG
'I was supposed to bake a cake today but I wasn't free'

(55) Xwla *Má nɔ̀ nɔ̀ ɔ̀xwé mɛ̀, vɔ̀ ùn tɔ́n.*
I-FUT already stay house in but I leave
'I should have stayed at home but I left'

Mu, *musu* and *ʍo* (*l/na*) are also employed to express probability (56).

(56) PM *A mu de a osu nounou.*
he must COP LOC house now
'He must be at home now'

Gen *Jan ʍó lá nɔ axómɛ fífijɛ́n.*
Jean must FUT COP house now.
'Jean must be at home right now'

Inferred certainty, however, is expressed by *musu* rather than *mu* in Pamaka and Ndyuka. In Saamaka and Gbe it is also conveyed by *musu* and *ʍo* (*l/na*), respectively. In

Gbe *мо* (*l/na*) is generally also combined with a phrase or adverb expressing certainty. Inferred certainty is exemplified in the examples in (57).

[you see a totally destroyed motorbike at a tree]

(57) PM *A man musu dede* (*tuutuu/ye*).
DET man must dead true-true/assertion
'The man must (surely) be dead' (Sanna, p.c. April 2003)

Waci *Kəkətɔ á мó lá kú kpò é.*
driver DET must FUT dead sure EMPH
'The driver must surely be dead' (Capo, p.c. Nov. 2003)

The literature (e.g. Aboh 2004; Lefebvre 1996, 1998; Lefebvre & Brousseau 2002) suggests that the Gbe varieties also have a subjunctive marker, *ni* in Fon and Gun, *ne* in Gen etc., see Table 2. It conveys a jussive sense, e.g. admonitions etc. (58a), or an optative sense, e.g. a wish or desire (58b).

(58) a. Gun *Kòfí ní jì hàn.*
Kofi INJ[19] sing song
'Kofi should sing a song.' (Aboh 2004: 181)

Aja *N ná gbe yi mɔ́ yi né vá kpɔ́m.*
I give voice her that she SUB come see-me
'I ordered her to come and see me'

Waci *Mə̀ gblɔ né-ɛ bə́ nə́ kplɔ xɔ mə.*
I tell to-her that INJ-she clean house in
'I told her to clean the house'

b. Fon *Máwu ní cɔ́ wè.*
God SUB protect 2sg
'May God protect you' (Anonymous 1983: V, 4 in Lefebvre & Brousseau 2002: 93)

Waci *Mə ji bə́ nə́ vá kpɔ́-m.*
I want/search that SUB-she come see-me
'I want her to come visit me'

It seems to also convey a hortative sense in at least some varieties (59), i.e. "the speaker is encouraging or inciting someone to action" (Bybee et al. 1994).

(59) Waci *Mí nə́ yi Kútɔnu.*
we SUB go Cotonou
'Let us go to Cotonou (lit. We should go to Cotonou)'

19. Aboh (2004) calls *ni* an injunctive marker when it occurs in main clauses and a subjunctive marker when it is used in subordinate clauses. Given the great similarity in meaning, I refer to both as subjunctive.

Unlike the Gbe varieties, the maroon creoles do not appear to have a subjunctive marker. The element *fu* or *fi* used in some of these functions in some Caribbean English Creoles (Winford 1985) may not precede the main verb of a clause to convey jussive (60) or optative meanings (61) in the Surinamese Creoles. The former sense is conveyed by *mu*(*su*) and the latter by *meki*.

(60) PM **A fu boli a nyanyan.*
he fu cook DET(sg) food
'He should cook the mean'
PM *A mu boli a nyanyan.*
he fu cook DET(sg) food
'He should/must cook the mean'

(61) PM **Gadu fu luku en.*
god fu look her
'May god protect her'
PM *Meki gadu luku en.*
make god look her
'May god protect her'

Fu is marginally acceptable in subordinate clauses though informants always indicate that *mu*(*su*) is preferred (62). It is more appropriately described as a preposition that may also function as a complementizer.

(62) PM *Mi taagi en taki a ?fu/mu boli a nyanyan.*
I tell him that he fu/must cook DET(sg) food
'I told him that he should/must cook the meal'
PM *Mi taagi en fu a boli a nyanyan.*
I tell him for he cook DET(sg) food
'I told him to cook the meal'

9. The categorial status of the modality elements

The Gbe elements that convey notions of necessity both appear to be auxiliaries because they cannot stand alone but always have to be followed by a main verb (63).

(63) Fon A: *Bàyí ní mà wɔ́ à?* B: **é ní.*
Bayi SUB prepare dough QP she SUB
'Must Bayi prepare dough? She must'
(Lefebvre & Brousseau 2002: 102)
Waci *Me ji a yi sukulu voa e mo la *(ji).*
NEG want FUT go school but he must FUT go
'He does not want to go to school but he must'

In the maroon Creoles, only *mu* in Ndyuka and Pamaka is a marker. As in the case of its Gbe counterparts, it cannot stand on its own but always has to be followed by the verb it modifies (64).

(64) PM A: *Mi mu kon*? B: **Iya i mu* *(kon).
I must come yes you must
'A: Do I have to come? B: Yes, you must'

In contrast to *mu*, *musu* seems to be a modal verb because its VP may be omitted (65). It either subcategorizes for a main verb (65) or for complement clause headed by the complementizer *fu* (51).[20]

(65) PM A: *Mi musu taki a toli de. B: Iya, i musu.*
I must talk DET(sg) story there yes you must
'A: I have to talk about this matter. B: Yes, you have to'

Abi 'have' is essentially a main verb that takes on modal meanings when it combines with *fu*.

10. The emergence of necessity in the creoles of Suriname

The investigation reveals close similarities between the maroon creoles of Suriname and Gbe in the expression of notions of necessity suggesting that Gbe influence also contributed to the emergence of this area of grammar in the creoles of Suriname.

Given the discussion above, it seems most likely that speakers of Gbe associated English phrases expressing strong obligation with equivalent native Gbe phrases and established an interlingual identity between English *must* and *ʍo* (*l/na*) in Gbe (66).

(66) English *They must eat.*
Gen *dɔkita ke le fiyé o, fada ke le fiyé o o ʍo*
doctor REL COP here PL priest REL COP here PL they must
la desi phla-gbe alo gɛngbe
FUT know xwla or gen
'The doctors and the priests who are [come] here, they must know [learn] Xwla or Gen' (Gen-NSF 3)

As a result of this association, the native speakers of Gbe projected the syntactic and semantic properties of their native Gbe element, *ʍo* (*l/na*), onto the English element, *mus*(*t*). In the early texts, for instance, *mus*(*t*), generally realized as *mo*, *moes*, *moesi* etc., functions as a marker of strong and weak obligation (67).

20. Note that *fu* is essentially a preposition but it may also function as a complementizer (cf. Winford 1985).

(67) SN *Joe mo krien drie pissi fossi befo …*
you must clean three piece first before
'You must first clean three pieces before …'
SN *Da dacteren takke joe moesi poeli mi bloede wan trom*
DET(sg) doctor talk you must pull my blood one time
morre.
more
'The doctor says you should bleed me once more'
(Arends & Perl 1995; van Dyk 1765)

Since *мо la* in Gbe also conveys weak obligation besides strong obligation (48)–(49) while English *must* only expresses strong obligation, it seems very likely that the use of *mu*(*su*) to convey weak obligation emerged exclusively due to Gbe influence. However, the other functions of *mu*(*su*) – probability, inferred certainty, and admonition – most likely emerged due to both influence from English and Gbe because these meanings are conveyed by both English *must* and Gbe *мо* (*l/na*).[21]

The differentiation between *mu* to mark strong and weak obligation and *musu* to convey inferred certainty in Pamaka and Ndyuka is most likely due to internal change. *Musu* probably shortened to *mu* in the contexts in which is was frequently used and involved little emphasis while the long form *musu* was retained in the less frequently used contexts that also involved emphasis. The use of *musu fu* as a marker of strong obligation is probably also due to an independent development in the creoles but at this point it is not clear which processes were involved. Finally, the marker of strong obligation *a*(*bi*) (*f*)*u* is most likely directly based on English *have to*.

11. Summary and conclusion

The discussion of the emergence of two subsystems of modality, potential and necessity, in the creoles of Suriname revealed that their structure in the modern creoles is the result of several kinds of processes. The comparative linguistic analysis revealed a number of close semantic and syntactic similarities between the modern maroon creoles and their main substrate, varieties of Gbe, suggesting that the latter played an important role in the emergence of these subsystems of grammar in the creoles of Suriname. Influence from the L1s of the main agents of creole formation cannot, however, account for all the properties of creole grammar since there are also differences between these creoles and the Gbe varieties. A consideration of the relevant superstrate sources (e.g. English, Dutch) showed very clearly that the European languages did not just provide the etymological shapes of creole functional elements but also contributed some of their semantic and syntactic properties. Some of the latter

21. It is not clear whether English *must* would have been used in all these functions in the plantation setting though.

features also coincided with substrate properties suggesting that the two main linguistic sources mutually reinforced each other. A consideration of the available early data suggests that the modality system of the creoles has undergone language-internal and contact-induced changes since its emergence. Some of the modal elements and certain properties of other elements in the modern creoles emerged gradually due to regular processes of language-internal change.

The findings clearly support a model of creole genesis that views creole formation as a multilayered process that shares many similarities with cases of L2 acquisition (cf. Siegel 1999; Winford 2002; Migge 2003). Given the nature of the contact setting (Arends 1995; Migge 2002, 2003), the creators of the plantation varieties came, for the most part, only into contact with (reduced) structures from European languages such as English, Dutch, and Portuguese. When they were able to establish interlingual identifications between European constructions and those in their L1s, they would adopt the former and partially or entirely reinterpret them according to L1 models. As a result of this interlingual association of structures, the ordering of elements and their etymological shape came to derive mainly from the superstrate varieties while the semantic and syntactic properties emerged either due to L1 influence alone or as the result of both superstrate and substrate influence, as in other cases of extreme L2 acquisition (Thomason & Kaufman 1988; Winford 2003). Once a subsystem of grammar had thus taken its initial shape, it became subject to other processes of contact-induced and language-internal change leading to the emergence of new modality elements (e.g. *man*) and changes in the semantic and syntactic properties of others (*mu*<*musu*).

The results from the investigation then clearly argue against a relexification account of creole genesis as proposed by Lefebvre (1996, 1998). First, the analysis suggests that the process of creole formation was not lexically but structure driven. The creators of the early creole 'targeted' and adopted entire structures rather than single lexical items from the superstrate varieties. The lexical elements that became part of the creole emerged from these structures due to the reinterpretation of these constructions according to L1 patterns and principles. Second, the L1s of the creators were only one of the sources for the properties of creole words and the subsystems of grammar they are part of. The syntactic and semantic features of creole items share important features with both their substrate and superstrate sources suggesting that the latter did not just contribute the etymological shapes of words but also played an important role in the emergence of their semantic and syntactic properties. Finally, the study also suggests that various processes of language-internal change, rather than just reanalysis, played a role in creole genesis.

References

Aboh, E. (2004). *The Morphosyntax of Complement – Head Sequences. Clause structure and word order patterns in Kwa*. New York NY: OUP.

Aboh, E. (to appear). Object shift, verb movement and verb reduplication. In G. Cinque & R. Kayne (Eds.), *Handbook of Syntax*. New York NY: OUP.

Arends, J. (1995). Demographic factors in the formation of Sranan. In J. Arends (Ed.), *The Early Stages of Creolization* (pp. 233–285). Amsterdam: John Benjamins.

Arends, J. & Perl, M. (1995). *Early Creole Texts. A collection of 18th-century Sranan and Saramaccan documents*. Frankfurt: Vervuert.

Berg van den, M. (2001). 'Mingo, joù no man'; Oud-Sranan in verhoren en verslagen van rechtszaken. *OSO, 20*, 241–253.

Berg, van den, M. Tense, mood and aspect in the early Sranan texts. Ms.

Bickerton, D. (1984). The Language Bioprogram Hypothesis. *The Behavioral and Brain Sciences, 7*, 173–221.

Bickerton, D. (1981). *Roots of Language*. Ann Arbor MI: Karoma.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The Evolution of Grammar. Tense, aspect, and modality in the languages of the world*. Chicago IL: University of Chicago Press.

Capo, H. B. C. (1988). *Renaissance du Gbe. Réflexions critiques et constructives sur l'Eve, le Fon, le Gen, l'Aja, le Gun, etc*. Hamburg: Helmut Buske.

Corne, C. (1983). Substratal reflections: The completive aspect and the distributive numerals in Isle de France Creole. *Te Reo, 26*, 65–80.

Dahl, Ö. (1985). *Tense and Aspect Systems*. Oxford: Blackwell.

Essegbey, J. (1999). *Inherent Complement Verbs Revisited. Towards an understanding of argument structure in Ewe* [MPI series]. Wageningen: Ponsen & Looijen.

Goury, L. (2003). Changements linguistiques dans le système de TMA des créoles du Surinam. Facteurs internes ou externes? presented at the *Journées d'Etudes Langues en contact* (*Fédération de Typologie des Universaux Linguistiques*).

Goury, L. & Migge, B. (forthcoming a). Between contact and internal development: Towards a multi-layered explanation for the development of the TMA system in the Creoles of Suriname. In S. Michaelis (Ed.), *Creoles between Substrates and Superstrates*. Amsterdam: John Benjamins.

Goury, L. & Migge, B. (forthcoming b). The mystery of *sa* in the Creoles of Suriname: Towards a gradual and multilayered account of Creole genesis. In R. Selbach, M. van den Berg, & H. Cardoso (Eds.), *Memorial Volume for Jacques Arends: Gradualist Approaches to Creole Genesis*. Amsterdam: John Benjamins.

Haeseryn, W., Romijn, K., Geerts, G., de Rooij, J., & van den Toorn, M. C. (1997). *Algemene Nederlandse Spraakkunst* (ANS) (2nd edition, 2 Vols.). Groningen/Deurne: Martinus Nijhoff/Wolters Plantyn. (http://oase.uci.kun.nl/~ans/)

Lefebvre, C. (1998). *Creole Genesis and the Acquisition of Grammar. The case of Haitian Creole*. Cambridge: CUP.

Lefebvre, C. (1996). The tense, mood and aspect system of Haitian Creole and the problem of transmission of grammar in creole genesis, *Journal of Pidgin and Creole Languages, 11*, 231–313.

Lefebvre, C. & Brousseau, A. M. (2002). *A Grammar of Frongbe*. Berlin: Mouton de Gruyter.

Migge, B. (1998a). Substrate Influence in the Formation of the Surinamese Plantation Creole: A consideration of sociohistorical data linguistic data from Ndyuka and Gbe. Ph.D dissertation, The Ohio State University.

Migge, B. (1998b). Substrate influence in creole formation: The origin of *give*-type serial verb constructions in the Surinamese Plantation Creole. *Journal of Pidgin and Creole Languages, 13*, 215–265.

Migge, B. (2000). The origin of property items in the Surinamese Plantation Creole. In J. H. McWhorter (Ed.), *Language Change and Language Contact in Pidgins and Creoles* [Creole Language Library 21] (pp. 201–234). Amsterdam: John Benjamins.

Migge, B. (2002). The origin of the copulas (*d/n*)*a* and *de* in the Eastern Maroon Creole. *Diachronica, 19*, 83–136.

Migge, B. (2003). *Creole Formation as Language Contact: The case of the Suriname Creoles* [Creole Language Libarary 25]. Amsterdam: John Benjamins.

Schumann, C. L. (1778). Saramaccanisch Deutsches Wörter-Buch. In H. Schuchardt (Ed.), *Die Sprache der Saramakkaneger in Surinam* 1914. Amsterdam: Johannes Müller.

Siegel, J. (1999). Transfer constraints and substrate influence in Melanesian Pidgin. *Journal of Pidgin and Creole Languages, 14*, 1–44.

Singler, J. V. (1990). *Pidgin and Creole Tense-Mood-Aspect Systems* [Creole Language Library 2]. Amsterdam: John Benjamins.

Singler, J. V. (1990). Introduction: Pidgins and Creoles and Tense-Mood-Aspect. In J. V. Singler (Ed.), *Pidgin and Creole Tense-Mood-Aspect Systems* [Creole Language Library 2] (pp. vii–xvi). Amsterdam: John Benjamins.

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 M. Harris & P. Ramat (Eds.), *Historical Development of Auxiliaries* (pp. 291–315). Berlin: Mouton de Gruyter.

Smith, N. (1987). The Genesis of the Creole Languages of Surinam. Ph.D dissertation, Universiteit van Amsterdam.

Smith, N. (2001). Voodoo Chile: Differential substrate effects in Saramaccan and Haitian. In N. Smith & T. Veenstra (Eds.), *Creolization and Language Contact* [Creole Language Library 23] (pp. 43–80). Amsterdam: John Benjamins.

Thomason, S. G. (1993). On identifying the sources of creole structures. In S. S. Mufwene (Ed.), *Africanisms in Afro-American Language Varieties* (pp. 280–295). Athens GA: University of Georgia Press.

Thomason, S. G. & Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley CA: University of California Press.

Winford, D. (1985). The syntax of fi complements in Caribbean English Creole. *Language, 61*(3), 588–624.

Winford, D. (2000a). Irrealis in Sranan: Mood and modality in a radical creole. *Journal of Pidgin and Creole Languages, 15*, 63–125.

Winford, D. (2000b). Tense and aspect in Sranan and the Creole Prototype. In J. H. McWhorter (Ed.), *Language Change and Language Contact in Pidgins and Creole* [Creole Language Library 21] (pp. 383–442). Amsterdam: John Benjamins.

Winford, D. (2002). Creoles in the context of contact linguistics. In G. Gilbert (Ed.), *Pidgin and Creole Linguistics in the Twenty-First Century* (pp. 287–354). New York NY: Peter Lang.

Winford, D. (2003). *An Introduction to Contact Linguistics*. Oxford: Blackwell.

Modeling Creole Genesis

Headedness in morphology

Tonjes Veenstra
John F. Kennedy Institute, Free University Berlin

In this paper we address two current claims in creole studies: (i) adults in the creolization are the innovators; (ii) the grammatical properties of creoles are either the result of continuities from the superstrate language(s) or from the substrate language(s), i.e. discontinuities are assumed not to exist. On the basis of a careful examination of the structure and emergence of synthetic compounds in the Surinamese creoles, in particular Saramaccan, we conclude that we cannot leave out the contribution of nativizing part of the population in the creolization process, as well as that discontinuities, i.e. properties that diverge from all the languages present in the original contact situation, do exist. As such, the research reported on here can be viewed as supporting the cascade-approach of DeGraff (1999).

0. Introduction

This paper addresses two current claims that have recently been gaining wide acceptance in creole studies.[1] The first claim take the adults in the creolization process to be the innovators, whereas the children are the regulators (Slobin 1977: 205). The generality of this claim will be severely questioned by showing that innovation also takes place during (or at) nativization. As such, the research reported on here can be viewed as supporting the cascade-approach of DeGraff (1999, 2002) in which "adults introduce innovative patterns into the linguistic ecology of language learners, whereas children, via L1A [i.e. processes of first language acquisition, TV], play a key role in restructuring adults' (*and their own*) innovations into stable grammars" (DeGraff 1999: 494; emphasis added [TV]). The basic point we want to make in this paper, therefore, is

1. Thanks to the editors and three anonymous reviewers for their insightful comments and constructive criticisms. I would like to acknowledge the work of Norval Smith on some of the topics discussed here. Usual disclaimers apply.

that we cannot leave out the contribution of the nativizing part of the population in the creolization process.[2]

A second, more implicit, claim is that one can solely account for the grammatical properties of creole languages by means of continuities from the superstrate language(s) or from the substrate language(s). As such, little room if any is left for discontinuities, i.e. those grammatical properties of creole languages that diverge from all the languages that were present in the original contact situation. The crucial question to ask, of course, is: do such discontinuities indeed exist? We show that they do.

This study focuses on the structure and emergence of synthetic compounds in the Surinamese creoles in general, illustrated with examples from Saramaccan (see Migge 2003; on Ndjuka & van den Berg 2003, on 18th century Sranan). Such compounds are formed by affixation of the suffix *-ma*. Examples are given in (1):

(1) a. ondosúku-tóngo-ma
research-language-ma
'linguist'

b. tjá-búka-ma
carry-mouth-ma
'messenger'

c. nái-koósu-ma
sew-clothes-ma
'tailor'

From a theoretical point of view, these constructions are of particular interest, since the affix is not attached to the head of the construction but rather seems to take a maximal projection of the head as the base to which it attaches. In other words, they constitute a case of phrasal affixation. As such, they are in violation of either Sadock's (1991) Strong Constructional Integrity or the Input Correspondence rule of Ackema and Neeleman (2000, 2002).

The paper is organized as follows. In Section 1 it is shown that the structure of the synthetic compounds in the Surinamese creoles is radically different from those

2. An anonymous reviewer points out that children seem to innovate less significantly than adults. Slobin (2002) and Tomasello (2002), for instance, give social reasons that make it difficult to credit children's creativity in the way we do in this paper. On the other hand, Henry and Tangney (1999), Roberts (1999), and Lightfoot (1999) as well as Rizzi (1999) report well-documented cases of 'grammatical invention' by children that are in line with the position adopted in this paper (see, e.g., DeGraff 1999: 509, 523f., 525 for discussion). Furthermore, with respect to the role of children in the emergence of Nicaraguan Sign Language (NSL), Senghas et al. (2004: 1780) observe that "research on NSL has found that changes in its grammar first appear among preadolescent signers, soon spreading to subsequent, younger learners, but not to adults". This results in a community in which the most fluent speakers are the youngest, most recent learners. It is our contention that a similar situation also existed in the history of Saramaccan (and creoles in general).

that are found in the languages that were present in the original contact situation. As such, they represent the 'discontinuity' referred to above. Furthermore, we show that *-ma* in contemporary Saramaccan has all the hallmarks of an affix. In Section 2, the leading question is to which process of language acquisition in the creole context the different word order inside synthetic compounds is due. We discuss the acquisition of such compounds from a SLA (second language acquisition) and a FLA (first language acquisition) perspective, and conclude that the difference in word order is due to FLA. From this, we conclude that innovation also takes place during nativization. We take this to be counter-evidence to the claim that the adults in the creolization process are the sole innovators, whereas the children are the regulators. Section 3 is devoted to an analysis of synthetic compounds. Section 4 presents a diachronic scenario for the emergence of synthetic compounds in the Surinamese creoles.

1. Affixation and synthetic compounds

The suffix *-ma* in Saramaccan is category-changing (the derivative being nominal in character) and it can attach both to nouns and verbs, as can be seen from the examples in (2):[3]

(2) a. édi-ma N-ma
head-ma
'boss, headcase'

b. sábi-ma V-ma
know-ma
'expert'

The examples in (2) are all instances of simple affixed words. In addition, however, the affix *-ma* can also be used to form synthetic compounds, in which it is combined with a complex constituent, superficially consisting of a noun and a verb, as in (3):

3. Veenstra (2006) discusses examples like (ia–b) in which it appears that *-ma* can also attach to adjectives and prepositions. However, due to the multifunctionality of lexical items in Saramaccan, an alternative analysis, as verb and noun respectively, can be given:

(i) a. giíi-ma adjective/(intransitive) verb
miserly-ma
'miser'

b. báka-ma preposition/noun
behind-ma
'afterbirth'

(3) a. bébe-daán-ma
drink-rum-ma
'drunkard'
b. séi-kónde-ma
sell-country-ma
'traitor'

In most (if not all) cases, the noun is interpreted as the object (internal argument) of the verb. We will argue in Section 3 that there is a verbal projection (VP) inside the synthetic compound.

As has been already recognized in the early nineties, e.g. Smith and Veenstra (1994), the verb object order in synthetic compounds presents a riddle, because the order in the creole (Saramaccan) is different from both the orders in the superstrate language and the major substrate languages. Whereas in Saramaccan the object follows the verb, it precedes the verb in English (the superstrate language) as well as the Gbe languages (the major substrate languages). This is exemplified in (4):

(4)	a.	truckdriver	*OV*	English
	b.	nàkí-sá-tó firewood-sell-AFF 'firewoodseller'	*OV*	Fon
	c.	téi-mánu-ma take-man-AFF 'maneater'	*VO*	Saramaccan

In addition to the reversed order of the verb and its object, there is another property that sets apart Saramaccan from both its super- and substrate languages. Whereas in Saramaccan the affix *-ma* is non-adjacent to the head of the synthetic compound, English as well as the Gbe languages have the affix adjacent to the head.

Still another difference relates to the form of the affix. Most of the Gbe languages have different affixes depending on the meaning (agentive vs attributive vs originative).[4] Although in English, the affix *-er* can be used in all these instances, other affixes, e.g. *-ian* in the originative, *-ist* in the agentive, also occur. Saramaccan, on the hand, uses the suffix *-ma* invariably. This is summarized in Table 1, and (partly) illustrated in (5):

(5)	a.	àzé-tò	witchcraft-AFF	'magician'	Fon
		hóndi-ma	hunt-AFF	'hunter'	Saramaccan
	b.	dèté-nò	deafness-AFF	'deaf person'	Fon
		tú-mujée-ma	two-woman-AFF	's.o. with two wives'	Saramaccan

4. Migge (2003: 82–83) observes that for some Gbe varieties (Gen, Waci, Xwela) one affix is *generally* used to denote this range of meanings. Her phrasing seems to imply that there are exceptions.

Table 1. Form of the affixes in possible substrates, superstrate, and creole*

	Fon	Vhe	Gen	English	Saramaccan	
Agentive	-tò	-(l)á	-tó	-er/-ist/Ø	-ma	(5a)
Attributive	-nò	-tó	-tó	-er	-ma	(5b)
Originative	-nù	-tó	-nò	-er/ian	-ma	(5c)

*Taken from Smith and Veenstra (1998)

c.	bènín-nù	Benin-AFF	's.o. from Benin'	Fon
	básu-sé-ma	bottom-side-AFF	's.o. living downstream'	Saramaccan

These differences between synthetic compounds in the creole and those in the super- and substrate languages are in our view sufficient to claim that we are dealing here with a discontinuity, i.e. a grammatical property of a creole language that diverges from all the languages that were present in the original contact situation. In other words, neither the superstrate language (English) nor the major substrate languages (Gbe) seem to have been the model for the verb-object order in synthetic compounds nor for the position of the affix. The interesting question, of course, is how the structure of these compounds came about in Saramaccan. Assuming a multidimensional model of creole genesis (cf. Corne 1994; DeGraff 1999; Kihm 2000; Muysken 2001; Becker & Veenstra 2003; Veenstra forthcoming, and below) in which processes of second as well as first language acquisition play substantial roles, the question reduces to to which factor in the creolization process this ordering difference is due: second or first language acquisition. Since the deviating word order is an innovation when compared to the structures of the contributing languages, the answer to this question can help us determine whether innovations are only instantiated by adults in the creole context or not.

This question is the focus of the next section, but first we turn to another issue. Until now, we have assumed but not actually shown that *-ma* is an affix. Smith and Veenstra (1990) present two arguments for the affixal status as *-ma* in contemporary Saramaccan, one phonological in nature, the other syntactic.

The phonological argument involves tone-polarization. Saramaccan, being a tone language, has two surface tones – high and low, and three underlying tones – high, low, and unspecified. In the analysis of Good (2004), *-ma* is analysed as having underlyingly a low tone. However, if the base to which it attaches, has a lexically-specified low tone in the final syllable, the low tone of *-ma* will be changed into a high tone:

(6) lègèdè-má
lie-AFF
'liar'

This makes *-ma* phonologically dependent and sets it apart from (independent) lexical items (free morphemes). A further argument is that it cannot occur on its own (the classic definition of an affix is that it is a bound morpheme, see Plag 2003 for a recent textbook treatment).

The second argument is syntactic in nature, and involves argument-inheritance. The argument runs as follows. If *-ma* is not an affix (but rather a lexical item), the synthetic compound would not be a derivational construct, but a genuine case of morphological compounding. A crucial difference between the two is that only in a derivational construct can the (internal) argument of the verbal base appear outside the construction (Di Sciullo & Williams 1987; Spencer 1991). The contrast is exemplified in (7)–(8) for English:

(7) a. [pizza-deliveryman] compounding
b. *[deliveryman] of pizzas

(8) a. [pizza-deliverer] derivation
b. [deliverer] of pizzas

As shown in (9)–(10), synthetic compounds in Saramaccan pattern with (8) and not with (7):

(9) a. [hóndi-fóu-ma]
hunt-bird-ma
'birdhunter'
b. [hóndi-ma] u dí fóu
hunt-ma FOR DET bird

(10) a. [súku-góuto-ma]
search-gold-ma
'golddigger'
b. [súku-ma] u dí góuto
search-ma FOR DET gold

The upshot is that synthetic compounds formed with *-ma* in Saramaccan pattern with synthetic compounds formed with *-er* in English. Since English *-er* has unequivocally an affixal status, we conclude that *-ma* is an affix as well.

Having settled the affixal status of *-ma*, we can now turn to the question which process of language acquisition in the creolization context is responsible for the difference in the order of the verb and its object in synthetic compounds.

2. Acquisition processes and synthetic compounds

While in the eighties and the nineties mono-dimensional models (e.g. language bioprogram hypothesis (Bickerton 1981, 1988) with an exclusive role for FLA, relexification hypothesis (Lefebvre 1998) with an exclusive role for SLA, etc.) were dominating the field of creole studies, at the turn of the 21th century multi-dimensional models appear to have gained ground (Corne 1994, 1999; DeGraff 1999, 2002, 2005; Kihm 2000; Mufwene 2001; Muysken 2001; Becker & Veenstra 2003; Veenstra forthcoming). As DeGraff (2003) so aptly observes, "the available linguistic and socio-historical details

and current results in linguistic-theoretical research, including language-acquisition research, seem to be more in line with a scenario in which both adult learners and child learners contributed to the 'creation' of Creole languages, each group in their own principled way".[5]

The issue we will deal with here is what the particular contribution of adult and child learners is. Two major processes are grammatical innovation and regularization, but is it the case, as already noted above, that innovation is solely restricted to the adult population in the creolization context? In order to arrive at an answer, we discuss the acquisition of synthetic compounds from different perspectives. Section 2.1. covers the SLA-perspective. Section 2.2. is on the FLA-perspective.

2.1 Second Language Acquisition

Broeder, Extra, van Hout and Voionmaa (1993), working within the Basic Variety-framework, observe that with respect to the acquisition of morphology in the first stages of SLA compounding precedes derivation. In fact, at this stage there is hardly any trace of productive derivation to be found.[6] They further note that initially there is clear evidence of source-language related influence in compounding.

As noted by Lardiere and Schwartz (1997), the acquisition of synthetic compounds is a relatively unexplored topic in SLA research. All the studies consulted for this paper (Boucher 1990 on French learners; Lardiere 1994, 1995; Lardiere & Schwartz 1997 on Spanish learners) report that 'errors' were made in affixation, verb-object order, and semantic interpretation, due to (UG-constrained) L1 influence. In what follows, we concentrate on the findings of Lardiere and Schwartz (1997).

---

5. As noted in Veenstra (2006), it might be countered that the socio-historical facts of Surinam do not provide evidence that there were enough children to play a significant role in forming creole structures. This point is driven home in its clearest formulation by Arends (1989, 1993, 1994) who remarks that for the Surinam colony during the first one hundred years in its existence, "no creolization – in the traditional view of nativization by children – could possibly have taken place simply because the children required to perform this process were not present" (Arends 1993: 375). Apart from the fact that we don't know how many children would be enough to introduce a creole language, Arends (1995) has now produced new figures which caused him to significantly revise his estimation upwards with respect to the number of children in the early history of the Surinam colony. From these new estimates Smith (2001: 51) concludes that "it is clear that we do not have much basis for claiming that at some earlier period there were too few children for the LBH to operate [in Surinam, TV]. The question that we must answer is rather, when could it have operated". We stand in full agreement with Smith's position.

6. Sporadically, an attempt was found in which a learner is trying out a derivational device, but overall such attempts were rare and certainly not characteristic of the early stage of SLA (Broeder, Extra, van Hout, & Voionmaa 1993: 56).

The study focuses on experimentally elicited data on the acquisition of English deverbal synthetic compounds by native speakers of Spanish. An example for each language is given in (11)–(12) respectively:

(11) lavaplatos — Spanish
lav -a plato -s
wash -3SG plate -PL 'dishwasher'

(12) body-snatcher — English

They note that synthetic compounds in these languages share at least three properties: (i) the deverbal nominalizations consist of a verbal head and its (internal) argument; (ii) the compound generally refers to a person performing the action of the predicate; (iii) they are highly productive in both languages. The three properties that set the languages apart are: (i) in Spanish the object is almost invariably pluralized, while in English plurals are regularly excluded in this position; (ii) in Spanish the verb carries 3SG agreement-features, but not in English. On the other hand, English has an overt nominalizing affix, but Spanish does not. However, both the 3SG agreement-features and the affix *-er* are spelled out on the (verbal) head of the construction; (iii) although in both languages the VP is head-initial in sentential syntax, only in Spanish is this canonical order retained. For the present discussion, the most interesting features are the order of the verb and its object and the position of the affix (recall that Saramaccan differs from its super- and substrate languages with respect to these features).

The experiment on which they base their results involved 34 adult native Spanish speakers (age 18–35) grouped by English L2 proficiency into three levels on the basis of the Michigan Test of English Proficiency (Form B): low (n=10), intermediate (n=12) and advanced (n=12). Overall, they found the following number of non-target responses (excluding pluralization errors):

Table 2. Rate of errorful compounds: all errors*

English level	% of total compounds produced
Low	71.13 (101/142)
Mid	43.09 (78/181)
High	04.21 (8/190)

*Adapted from Lardiere and Schwartz (1997)

As can be seen from Table 2, proficiency level and the rate of errorful compounds are inversely correlated. In their paper, Lardiere and Schwartz (1997) discuss two types of non-target responses: VO-error and ING-error.

The most common type of nontarget response involves compounds which exhibit inverted word order, namely, VO. Some proto-typical examples are given in (13):

(13) 'eater-flies', 'washing-hand', 'kissing-baby mother'
(as opposed to 'fly-eater', 'hand-washer', 'baby-kissing mother')

Table 3 shows that the low-proficiency group shows the most variable behaviour (half of the compounds show the inverted order), and that the VO-error has almost completely disappeared at the high proficiency level:

Table 3. Rate of VO order inside compounds*

English level	% of total compounds	% of errorfully produced compounds
Low	41.55 (59/142)	58.42 (59/101)
Mid	22.65 (41/181)	52.56 (41/78)
High	00.53 (1/190)	12.50 (1/8)

*Adapted from Lardiere and Schwartz (1997)

The second type of non-target response involves compounds in which the verbal suffix *-ing* appears in contexts obligatorily requiring the (agentive) *-er* suffix instead, as exemplified in (14):

(14) 'washing-hand(s)'
(Instead of hand(s)-washing; when not accompanied by an overt agent)

As can be read from Table 4, the ING-error displays a similar acquisitional development as the VO-error, i.e. they were mainly produced in the two less proficient groups of subjects:

Table 4. Rate of ING error*

English level	% of total compounds	% of errorfully produced compounds
Low	10.56 (15/142)	14.85 (15/101)
Mid	14.92 (27/181)	34.62 (27/78)
High	01.05 (2/190)	25.00 (2/8)

*Adapted from Lardiere and Schwartz (1997)

The analysis Lardiere and Schwartz (1997) propose crucially assumes it is the L1 grammar that is responsible for the non-target responses. The VO order in the English synthetic compounds is a direct reflection of the Spanish word order in such compounds. They argue that the interlanguage grammar underlyingly still makes use of the 3SG features (positioned on the verb in Spanish), and that these features get spelled out either as the affix *-er* or the affix *-ing*. Thus, the order of the verb and its object and the position of the affix are due to (UG-constrained) L1 influence.

Finally, they note that another, perfectly feasible error is virtually nonexistent in their data. It is not the case that the learners produce compounds like *chase-micer* where the *-er* is attached to the entire VP rather than just to the verb, even though they hear *-er* as the last (linear) element of the compound in the input.

Turning now to the Saramaccan case, it should be clear that if the verb-object order as well as the position of the affix in interlanguages is due to L1 influence in SLA,

the expectation is that the order of the Gbe languages (NV) would have survived in Saramaccan, contrary to fact. Also, we would expect the affix to be generated on the (verbal) head, again contrary to fact. Interestingly, the pattern we find in Saramaccan is exactly the pattern that the learners studied by Lardiere and Schwartz did not produce. From this we conclude that processes of SLA were not responsible for the innovative patterns that we find in synthetic compounds in Saramaccan.

2.2 First Language Acquisition

Studies on the L1 acquisition of synthetic compounds (Clark & Hecht 1982; Clark, Hecht, & Mulford 1986; Clark & Berman 1987; Clark & Barron 1988; Clark 1993) show that initially children have problems in determining the correct position of the affix as well as the order of the verb and its object. The latter is particularly unstable in the first few years (in production as well as comprehension). For instance, Clark, Hecht and Mulford (1986) conducted a study in which they subjected 48 children ranging in age from 3;0 to 6;11 to an elicitation experiment. Their results can be summarized as follows:

(15) a. **Ordering Errors**. The youngest children used incorrect order more often than the older groups. *V-ing + O, *V-er + O and *V + O accounted for 24% of compounds elicited from the 3;6 group compared to only about 2% of the 6;5 group.

b. **Affixation Errors**. Approximately 10% of the 3;6 age group's coinages were bare stem compounds (*V + O). In contrast, the older groups made virtually no affixation errors. Few instances of affixation on the Object constituent were elicited.

Three stages can be distinguished. By the first stage, children are able to produce N + N, agentive V + MAN (e.g. *fix-man*) and instrumental V + THING (e.g. *cut-machine*) compounds. At this point, the verb and its object appear in *compound order*. Clark, Hecht and Mulford (1986) suggest that children have made a generalization about head position: the rightmost noun designates the semantic category (perhaps by analogy, since their lexicon presumably contains similar forms such as *milkman* and *fireman*).

The second stage can actually be split up in two, the availability of the affix *-er* being the watershed. As shown by Clark, Hecht and Mulford (1986), children use the VO order typical of English syntax in such compounds before they start using the affix *-er*:

(16) *Stage IIa* (*around age 3*)*: VO order, no overt affix*

a. a kick-ball (someone who kicks a ball)
b. a build-wall (someone who builds a wall)
c. a bounce-ball (someone who bounces a ball)

They note, "essentially, what children at this stage appear to do is nominalize the VERB PHRASES in the descriptions they hear" (Clark, Hecht, & Mulford 1986: 22).

They discuss a subsequent acquisitional stage in which the affix *-er* is spelled out, but the verb and its object still show up in the head-first order typical of English syntax:

(17) *Stage IIb* (*around age 4*)*: VO order, overt affix on either V, N or both* (*in order of frequency*)

a.	a giver-present	(someone who gives a present)
b.	a dry-hairer	(someone who dries hair)
c.	a mover-boxer	(someone who moves boxes)

The affix is in the majority of cases rightly attached to the (verbal) head of the construction. In these cases, however, the affix *-er* still has wide scope at Logical Form (LF) (Randall 1982; Roeper 1997):

(18) a. kicker-ball
LF: (kick ball)-er
b. There's a bike-rider with no hands.
LF: (ride bike with no hands)-er

The third stage consists of the acquisition of N-V compounding, with the consequent possibility of synthetic compounding. Interestingly, correct ordering of the verb and its object was always produced in conjunction with correct placement of the deverbal suffix; there were no elicitations of bare stems in O + V order. Clark, Hecht and Mulford (1986) argue that the absence of this pattern indicates that children learn to apply affixation to the verbal stem before learning correct constituent order.

Summarizing, we can say that the verb-object order in FLA is particularly unstable, and that children start out with the syntactic head-initial order of the (verbal) head and its (internal) argument.[7] The verb-object order in Saramaccan corresponds to the orders at stage 1 and 2. As a likely explanation, we suggest that the child learners in the creolization context stuck to this order due to the structurally-underdetermined

7. A potential problem, depending on one's analysis of West Germanic verbal syntax, either as head-final (Koster 1975) or head-initial (Zwart 1993), might appear to be the fact that Dutch and German children also (occasionally) make similar ordering errors:

(i)	a.	openflessener	(open-bottle-er)	'bottle-opener'	(Verrips 1998)
		wisseruiters	(wipe-glass-er)	'windshield-wiper'	
	b.	springrunter	(jump-down-er)	'water coming out of a bottle'	(Vincent, 2;7)

The problem would disappear if it can be shown that these structures are derived CPs (with the verb in C) rather than VPs, since in anyone's analysis CPs are head-initial. In any case, the data show that the instability of verb-object order is not a particular quirk in the acquisition of English. A full treatment is left to future research.

(pidgin-) input they encountered.[8] The crux of this explanation is that, as we have seen above, there is a correlation between the head-final order and the presence of the affix, i.e. without the affix no head-final order. It is reasonable to assume that the English affix *-er* did not survive the pidginization process (as well as the other affixes of English for that matter in the Surinamese creoles). So, there was no evidence for the child learners to bootstrap themselves into stage 3 in which the head-final order kicks in. Therefore, they stuck to the head-inital order. The conclusion, then, is that the verb-object order in synthetic compounds is due to FLA.

2.3 Summary of implications

We started out with the issue of what the particular contribution of adult and child learners in the creolization context is. Two major processes are grammatical innovation and regularization, and the question was whether innovation is solely restricted to the adult population in the creolization context.

We identified two innovative features of synthetic compounds: (i) head-initial word order; (ii) non-adjacent position of the affix. We argued that if the verb-object order as well as the position of the affix in interlanguages is due to L1 influence in SLA (as shown by Lardiere & Schwartz 1997), the order of the Gbe languages (NV) would have survived in Saramaccan and the affix would be generated on the (verbal) head, both contrary to fact. This rules out that these innovative features are due to processes of SLA. Contrarily, we argued that the inverted order of the verb and its object is due to the structurally-underdetermined input the child learners in the creolization context were confronted with. As such, processes of FLA are ultimately responsible for introducing this innovation. Thus, innovation also takes place during nativization, and is not solely limited to the adult population in the creolization context.

This leaves us with the other innovative feature. The position of *-ma* in synthetic compounds, if it were to have affixal status, is not easily accounted for in the scenario sketched above. The reason is that in both SLA and FLA learners do attach the affix on

8. As an astute reviewer remarks, it seems necessary to determine the extent to which adults' interlanguages are 'structurally-underdetermined' and whether the extent and the ways in which they might be underdetermined match the morphological profile of the 'pidgin' input that seeded Saramaccan's synthetic compounds. From the available literature, it is clear that derivational morphology (for both agent and non-agent reference to entities) only plays a minor role in early stages of SLA. Thus, Broeder, Extra, van Hout and Voionmaa (1993: 70) observe that (i) suffix types are mostly represented by only one form per informant; (ii) several representatives are direct imitations of the native interlocutor; (iii) various suffix types are in fact lexicalisations or formulas. That is, the use of derivational morphology is extremely limited and, if derivational affixes are present, they show random variation and are not used innovatively. Assuming that 'pidgins' can be equated with 'adults' early interlanguages" (e.g. DeGraff 1999: 494, but the assumption can be traced back to (at least) Bickerton 1977), we conclude that the pidgin-input in Suriname was structurally underdetermined with respect to the realm of derivational morphology. In fact, there is no trace of any English suffix to be found in the Surinamese creoles.

the (verbal) head of the compound from early on. A possible way out of this deadlock is to propose that *-ma* was not at all an affix in the formative years of Saramaccan. We spell out this proposal in Section 4.

The introduction of these two innovations into the domain of synthetic compounding in Saramaccan has led to a cross-linguistically rare configuration that poses interesting theoretical questions, an issue we turn to now.

3. Analysis

The general model of grammar we assume is that of Representational Modularity (Jackendoff 1997, 2002), in which phonology, semantics and syntax are independent generative systems associated by mapping principles. In this framework, phonological and syntactic representations are not isomorphic, and phonological and syntactic primitives are members of disjoint sets, as exemplified for Saramaccan in (19):

(19) a. I festan no?
2SG understand Q
'Do you understand?'

b. $[_{CP}\ [_{IP}\ [_{DP}\ \text{i}]\ [_{VP}\ \text{festan}]]\ \text{no}]$ syntax

c. $[_{\varphi}\ [_{\omega}\ \textbf{i}]\ [_{\omega}\ \textbf{fɛ sta}\upsilon]$ phonology

Turning to morphology, affixes as bound morphemes are traditionally seen as monolithic entities, i.e. an affix selects for a head of a certain category. Thus, Ackema and Neeleman (2002) observe that affixes contain both (morpho-)syntactic (the category of the base) and (morpho-)phonological information (the base is a head). Whereas the syntactic information can differ from affix to affix, the phonological information invariably holds for all affixes (it is thus understood as one of the defining properties of what an affix is). Also in the Chomskyan tradition (e.g. Chomsky 1993), it is usually assumed that these two types of information are indivisible properties of a single element (the affix), which is 'inserted' (in GB-terms) or 'merged' (in minimalist terms) in the syntax.

There is an alternative, however. In the 'separationist' view (cf. Sproat 1985; Anderson 1992; Halle & Marantz 1993; Beard 1995), no joint insertion of syntactic and phonological information takes place. Instead, syntactic features are only present in syntax, whereas phonological features are only present in phonology, and the two are related by correspondence rules. In other words, a lexical item (bound or free morpheme) is nothing but a minimal, idiosyncratic, version of such a rule. Note that this conception is highly compatible with Representational Modularity.

Ackema and Neeleman (2000, 2002) propose it is not only the representation of words and phrases themselves, but also the representation of their selectional properties that must be distributed over the different modules. We will follow their proposal as well as their notation. From now on, the syntactic information of bound morphemes is represented as AFFIX, while the phonological information is represented

as /affix/. In this sense, what we have been calling an affix up to now is actually a (minimal and idiosyncratic) rule mapping a particular AFFIX on a particular /affix/.

In addition to idiosyncratic mapping rules, there are also general mapping rules (or principles) that regulate the interaction between the syntactic and the phonological components. The mapping between the syntactic and phonological structure of words is restricted by four mapping principles. These rules are in general non-violable.[9] The first mapping principle can informally be stated as saying that there should not be crossing correspondences between morphosyntactic and morphophonological structures:

(20) *Linear Correspondence*
If X is structurally external to Y,
X is phonologically realized as /x/, and
Y is phonologically realized as /y/
then /x/ is linearly external to /y/.

The second mapping principle says that phonological selectional features cannot be satisfied by just any adjacent word; they must be satisfied by the correspondent of the AFFIX's input:

(21) *Input Correspondence*
If an AFFIX selects (a category headed by) X,
the AFFIX is phonologically realized as /affix/, and
X is phonologically realized as /x/,
then /affix/ takes /x/ as its input.

The third mapping principle prevents multiple phonological realizations of a single affix (just like the uniqueness principle of Noyer 1993):

(22) *Quantitative Correspondence*
No element in the morpho-syntax is spelled out more than once.

The last mapping principle ensures that the AFFIX is phonologically realized if a phonological realization is listed for the AFFIX in the lexicon:

9. To be more specific, regular mappings can be overruled on an idiosyncratic basis. That is, general mapping principles seem to be 'elsewhere conditions' that must be obeyed unless more specific rules apply. This means that violations of mapping principles are different from violations of syntactic or phonological constraints: whereas the latter invariably lead to ungrammaticality, otherwise illegal mappings can be grammatical if the lexicon contains a more specific rule. See Ackema and Neeleman (2000, 2002) for detailed discussion of possible violations and examples thereof.

(23) *Lexical Faithfulness*

If X is associated with /x/ in the lexicon, and
X occurs in the morphosyntactic representation
then /x/ must occur in the morphophonological representation.

Three of these principles (Linear Correspondence, Quantitative Correspondence, and Lexical Faithfulness) are general constraints that require parts of a complex word to be spelled out once in the appropriate linear position. The fourth principle (Input Correspondence in (21)) is relevant to structures involving affixation. It is this principle that will play an important role in the remainder of this section, as it dictates to which word an /affix/ should be attached.

Turning now to Saramaccan, we have presented clear phonological and syntactic evidence in Section 1 that *-ma* is an affix in Saramaccan.[10] The mapping rule for *-ma* would then amount to the following (based on the mapping rule for *-er* in Ackema & Neeleman 2002):

(24)

<=>x [x PRED (x)]	<=> [<–V, +N>, ($\underline{R}_x$)]	<=> /-ma/
	+[<±V, ±N> __]	+[$_\omega$ /y/ __]
SEMANTICS	SYNTAX	PHONOLOGY

Thus, *-ma* can be characterized as an affix from a syntactic (AFFIX) as well as a phonological (/affix/) point of view. Still, as we have seen, *-ma* differs from the run-of-the-mill affixes in that it can attach to expanded constituents:

(25) [púu-tánda]-ma
pull-tooth-ma
'dentist'

The question is whether the constituent *[púu-tánda]* is a morphological or a syntactic object. We can rule out a morphological analysis on two grounds. First, (almost) all compounds in Saramaccan are right-headed:

(26) a. tjúma-goón
burn-land
'burned land'

b. dóki-físi
dive-fish
'recently-caught fish'

Second, the noun can be modified by an adjective. Therefore, a maximal projection (AP) is present inside the constituent. Since neither determiners nor numerals (which we analyze as functional heads in the nominal extended projection) are possible, we

10. This rules out an alternative analysis, in which *-ma* would be analysed phonologically as a /word/. Thanks to Ingo Plag for pointing this out.

are dealing with a NP here (rather than a DP or NumP).[11] This rules out a strictly morphological analysis:

(27) a. tíi-{*dí, *déé, *tu, gaán}-ópoláni-ma
steer-{DET, DET, two, big}-aeroplane-ma
'pilot of big planes'
b. $[_{N}\ [_{VP}$ V {NP/*DP}] -ma]

The conclusion we draw is that we are dealing with a syntactic object, in particular a VP. As such, synthetic compounds violate *Input Correspondence*-mapping rule, as we will show now.

Within the framework of Ackema and Neeleman (2000, 2002) AFFIXes in syntax can take a phrase as their host, but it is the phonological counterpart, the /affix/, that selects a word. By *Input Correspondence* of (21), the word to which an /affix/ attaches corresponds to the category selected by the AFFIX, *or to its head if the selected category is a phrase*. In general, then, the following configuration is ruled out:

(28) a. $[_{Y}\ [_{XP}$ X WP] AFFIX] syntax
b. */x/-/wp/-/affix/ phonology

If associated with (28a), (28b) violates *Input Correspondence*, not any of the other correspondence rules. On the basis of the evidence in (27), we conclude that synthetic compounds in Saramaccan have the configuration in (28), and, therefore, violate this mapping principle.

As Ackema and Neeleman (2002) point out, none of the mapping principles are violated if prior to the attachment of the overt affix a process of zero derivation has taken place. In other words, the syntactic structure in (29a) can successfully be mapped onto the phonological one in (29b).

(29) a. $[_{Z}\ [_{Y}\ [_{XP}$ X WP] AFFIX_1] AFFIX_2]
b. /x/-/wp/-/affix_2/

Due to *Input Correspondence*, /affix_2/ must attach to the phonological correspondent of AFFIX_1. However, since AFFIX_1 does not have a correspondent, *Input Correspondence* will be satisfied vacuously. Therefore, /affix_2/ is allowed to freely attach to any adjacent word.

Evidence that *-ma* belongs phonologically to the adjacent word, yet from a syntactic (and semantic) point of view takes the whole phrase as its base, is given in (30):

11. We follow here the proposal of Sportiche (1999) that the extended projection associated with the object is merged external to the VP. Sportiche adduces evidence for his proposal from the lack of reconstruction of Determiners into the VP, and the Conservativity of Determiners, as argued for by Keenan (1981), as well as the fact that cross-linguistically DPs are excluded from (synthetic) compounds.

(30) hóndi-lògòsò-má.
hunt-turtle-ma
'turtle-hunter'

-*Ma* receives a high tone due to the adjacent low tone from the noun lògòsò 'turtle' (cf. the discussion of (6) with respect to tone-polarization in Section 1 above).

This analysis does not violate any of the proposed mapping principles regulating the syntactic and phonological structure of words (neither *Linear Correspondence, Input Correspondence*, *Quantitative Correspondence* nor *Lexical Faithfulness*).[12] The attachment of the affix -*ma* to a maximal projection is mediated through zero-affixation (i.e. an AFFIX without a corresponding /affix/). This analysis also confirms one of the typological predictions of Ackema and Neeleman (2002: 37): "mixed categories (...) must involve zero affixes in case the syntax is head-initial and the morphology characterized by the right-hand head rule". Needless to say, Saramaccan displays these typological properties.

The function of the zero-affix is that it nominalizes the VP. Nominalization processes in Saramaccan are without exception due to zero-affixation:

(31) [$_{NP}$Júu móni] tá tjá toóbi kó a i.
lend money ASP carry quarrel come P 2SG
'Lending money gets you in a quarrel.'

This means that we need a process of zero-affixation to account for the grammar of Saramaccan independently, thus taking away a possible objection that zero-affixation is introduced into the analysis merely to save the structure of synthetic compounds from violating *Input Correspondence* (which would otherwise be *ad hoc*).[13]

---

12. Veenstra (2006) discusses and rejects an alternative analysis in which synthetic compounds actually *do* violate *Input Correspondence*. The main problem with this analysis is that whereas in the cases of the English children and adults there is variation in which mapping principle can be violated in order to form such compounds, in the Saramaccan case violation of *Input Correspondence* is not only be the norm, but also categorial. Ackema and Neeleman (2002: 45) argue that only derived compounds that do *not* violate any of the mapping principles block other possible realizations.

13. Veenstra (2006) concludes that in this analysis it is not the affix -*ma* that nominalises its base (that's the function of the zero affix). The syntactic representation of -*ma* in (24) has two tasks: (i) it imposes (morpho-)syntactic selectional requirements, i.e. it selects for a nominal or a verbal category; (ii) it determines the categorial status of the whole construction. Due to the multifunctionality of lexical items in Saramaccan, i.e. all verbs, except the copulas *dé* and *dá*, also occur as nouns, it may be even argued that -*ma* only selects a nominal category.

4. The categorial status and position of -MA and diachronic depth

In Section 1 it was shown that in contemporary Saramaccan *-ma* has all the hallmarks of an affix. In Section 2 it was shown that once learners identify an element as an affix, they (almost) invariably attach it to the head of the synthetic compound. This seems to be true for both adult and child learners, although some variation is found with the latter. Still, the position of *-ma* in synthetic compounds in Saramaccan is not adjacent to the (verbal) head. This would be quite unexpected had *-ma* an affixal status right from the beginning. We, therefore, hypothesize that in the formative years *-ma* started out as a lexical word.

The idea that the suffix *-ma* is the result of a morphological change dates back to Arends (1995), based on historical sources of Sranan (he argued for this change on the basis of Sranan form *-man*). His main argument is "the fact that *man* is used to refer to females as well as to males (e.g. *helpiman* lit. 'help-man' for 'midwife', Schumann p. 128) is the result of a grammaticalization process, which made possible the use of *man* as a gender-neutral agentive suffix" (Arends 1995: 42). He claims that in the earliest source (Van Dyk, c 1765) primarily circumlocutions are used (e.g. *wan somma dissi loeke schribi zanti* lit. 'a person which looks written things' for 'a clerk'). Later sources (Schumann 1783 & Wullschlägel 1856) have a preference for compounding with *man* to express the concepts described by Van Dyk (c 1756) in terms of circumlocutions. Recently, this approach has been questioned by Braun (2001) and van den Berg (2003).

Braun (2001: 53–54) challenges Arends' position by showing that in Van Dyk quite a number of complex words with *-man* are attested (the exact number is 23 which is quite a lot, given the comparatively small size of the source). Moreover, the status of circumlocutions in Van Dyk is unclear, since in some cases both a circumlocution and a *-man* word are provided as references to the same occupation so that some circumlocutions can be analyzed as explanations provided by Van Dyk. Thus, from the fact that Van Dyk contains a number of circumlocutions denoting occupations it does not necessarily follow that the preference for compounding patterns is a later development.

Van den Berg (2003) shows that even in Van Dyk (c 1756) gender-neutral uses of *man* are attested. On the basis of this gender-neutral usage, she, as well as Bruyn (2002) and Migge (2003), argues that the form *man* was a suffix from the start. We will not follow their argumentation for the following reason: although there is no clear evidence that *-ma* was not an affix in the formative years, there is also no clear evidence to the contrary, because semantic extension (i.e. gender-neutrality) is just not indicative of morpho-syntactic status. For example, from the fact that in Dutch *postbode* 'postman' can refer both to males and females, it does not follow that *bode* 'courier' is a suffix (which it is not). Similarly, in German the words *das Mädchen* 'the girl' and *der Kerl* 'the guy' can be used in a gender-neutral sense, as in (32):

(32) a. Er ist ein Mädchen für Alles. German
'He is a handyman'

b. Sie ist ein lieber Kerl.
'She has her heart in the right place/She's a nice guy'

This does not mean, however, that these items have lost their morpho-syntactic status of independent words. Therefore, the verdict on the grammatical status of *-ma* in the formative years is still open.

However, we do maintain that the non-adjacent position of *-ma* in the synthetic compound is a strong argument for its non-affixal usage early in its history. Unfortunately, synthetic compounds are completely absent in the early sources of the Surinamese creoles, so we can only assume that non-adjacency has always been the case.[14]

We propose the following scenario. As noted in Section 2.1 in the earliest stages of SLA compounding precedes derivation. We, therefore, assume that adult learners in the creolization context also primarily used compounding strategies, and formed simple compounds with *ma*(*n*).[15] Also, the child learners first form simple compounds (see the discussion of stage 1 in Section 2.2 above). The behaviour of these two groups of learners is similar in this respect, and reinforces compounding as the main (or even only) strategy for forming complex words in the initial stages. The older form (*man*) was thus phonologically not an affix, but a word(/stem). Since in compounding there is basically no subcategorization, *man* was able to attach to all kinds of heads and phrases. Later on in the diachrony of Saramaccan, it has been grammaticalized into an /affix/. This is represented schematically in (33):

(33) Development in Saramaccan

Compounding	*ma*(*n*)	synt: WORD
		phon: /word/
intermediate	*ma*	synt: AFFIX
		phon: /word/
Affixation	*ma*	synt: AFFIX
		phon: /affix/

14. Van den Berg (2003) notes there is only one construction resembling a synthetic compound: *boonjamman* lit. 'bone-eat-man' for 'someone who suffers from gout or rheumatism'. Interestingly, the verb-object order is the opposite of what we find in the Surinamese creoles today. However, since *boonjam* is listed separately as 'gout, rheumatism', van den Berg (2003) argues that this is not an instance of a synthetic compound.

15. This is in line with the observations of two anonymous reviewers that in all varieties of English, whether vernacular or not, whether standard or not, compounds such as *milkman, garbageman, waterman, postman, fireman, mailman*, etc are present, which must have provided the initial input. As one of them notes, the model seems so productive that in Jamaica, for instance, Asians are called *Chinaman*, suggesting that what has developed in Saramaccan is not so isolated. A full treatment of compounds in other English-related creoles is beyond the scope of the paper.

The innovation in the last phase is the introduction of zero-derivation. In terms of economy, adding a zero-affix to your grammar may be seen as less costly than changes in the linearization of the elements in the compound which are PF-visible.

The advantage of this scenario (over one without diachronic depth) is that it can account for the fact that *-ma* ends up on the right edge, the reason being that child learners from early on already make a generalization about head position in (simple) compounds: the rightmost noun designates the semantic category. Furthermore, (almost) all compounds in Saramaccan are right-headed (see Section 3). We can also explain the verb-object order. Since the acquisition of the affix is a prerequisite for establishing the NV-order (see Section 2.2), this pattern could not have emerged precisely because of the absence of an affix-like element in the early stages.

Contrastively, in a scenario without diachronic depth it remains a mystery as to why the affix occurs at the right edge, because of the following two reasons: (i) the SL-learners of Lardiere and Schwartz (1997) do not produce such compounds (see Section 2.1); (ii) both child and adult learners in the majority of cases do attach the affix onto the head of the synthetic compound. The observed verb-object order seems to be even more of an embarrassment to this scenario.

The scenario we propose has the following implications: (i) there exists a discontinuity between the sub- and superstrate languages in the original contact situation and the resulting contact language. In other words, we cannot derive all the properties of creole languages from the properties of the languages that were present in the contact situation; (ii) this discontinuity is the result of a breakdown of grammatical cohesiveness (i.e. compounding strategies replace affixal ones); (iii) it is not the case that innovation only takes place in the adult population. The cross-linguistically rare properties of synthetic compounding in Saramaccan are due to innovations within the child population in the creolization context of Suriname.

References

Ackema, P. & Neeleman, A. (2000). M-selection and phrasal affixation. *UCL Working Papers, 12*, 307–344.

Ackema, P. & Neeleman, A. (2002). Morphological selection and representational modularity. In G. Booy & J. van Marle (Eds.), *Yearbook of Morphology 2001* (pp. 1–52). Dordrecht: Kluwer.

Anderson, S. (1992). *A-Morphous Morphology*. Cambridge: CUP.

Arends, J. (1989). Syntactic Developments in Sranan. Ph.D dissertation. University of Nijmegen.

Arends, J. (1993). Towards a gradualist model of creolization. In F. Byrne & J. Holm (Eds.), *Atlantic Meets Pacific* (pp. 371–380). Amsterdam: John Benjamins.

Arends, J. (1994). The African-born slave child and creolization (a postscript to the Bickerton-Singler debate on nativization). *Journal of Pidgin and Creole Languages, 9*, 115–119.

Arends, J. (1995). The Sranan Texts. In J. Arends & M. Perl (Eds.), *Early Surinam Creole Texts. A collection of 18-thcentury Sranan and Saramaccan texts* (pp. 11–93). Frankfurt: Vervuert.

Beard, R. (1995). *Lexeme-Morpheme Base Morphology*. Albany NY: SUNY Press.

Becker, A. & Veenstra, T. (2003). The survival of inflectional morphology in French-related creoles: The role of SLA processes. *Studies in Second Language Acquisition, 25*, 283–306.

Bickerton, D. (1977). Pidginization and creolization: Language acquisition and language universals. In A. Valdman (Ed.), *Pidgin & Creole Linguistics* (pp. 49–69). Bloomington IN: Indiana University Press.

Bickerton, D. (1981). *Roots of Language*. Ann Arbor MI: Karoma.

Bickerton, D. (1988). Creole languages and the bioprogram. In F. Newmeyer (Ed.), *Linguistics: The Cambridge Survey* (pp. 267–284). Cambridge: CUP.

Boucher, P. (1990). Teaching compound nouns in English. *CIEREC/TRAVAUX LXXVI*, Université de Saint Etienne.

Braun, M. (2001). Complex Words in Early Sranan: An investigation into creole morphology. MA thesis, Universität Siegen.

Broeder, P., Extra, G., van Hout, R., & Voionmaa, K. (1993). Word formation processes in talking about entities. In C. Perdue (Ed.), *Adult Language Acquisition: Cross-linguistic perspectives* (pp. 41–72). Cambridge: CUP.

Bruyn, A. (2002). The structure of the Surinamese creoles. In E. Carlin & J. Arends (Eds.), *Atlas of the Languages of Suriname* (pp. 153–182). Leiden: KITLV Press.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The View from Building 20* (pp. 1–52). Cambridge MA: The MIT Press.

Clark, E. (1993). *The Lexicon in Acquisition*. Cambridge: CUP.

Clark, E. & Hecht, B. (1982). Learning to coin agent and instrument nouns. *Cognition, 12*, 1–24.

Clark, E. & Barron, B. J. S. (1988). A thrower-button or a button-thrower: Children's judgements of grammatical and ungrammatical compound nouns. *Linguistics, 26*, 3–19.

Clark, E. & Berman, R. (1987). Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child Language, 14*, 547–567.

Clark, E., Hecht, B., & Mulford, R. (1986). Coining complex compounds in English: Affixes and word order in acquisition. *Linguistics, 24*, 7–29.

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. (1999). *From French to Creole*. London: University of Westminster Press.

DeGraff, M. (1999). Creolization, language change and language acquisition: An epilogue. In M. DeGraff (Ed.), *Language Change and Creation: Creolization, diachrony and development* (pp. 473–543). Cambridge MA: The MIT Press.

DeGraff, M. (2002). Relexification: A reevaluation. *Linguistic Anthropology, 44*, 321–414.

DeGraff, M. (2003). On Creole Genesis and Language Acquisition: Some methodological and theoretical preliminaries. Ms., MIT.

DeGraff, M. (2005). Morphology & word order in 'creolization' & beyond. In R. Kayne & G. Cinque (Eds.), *Handbook of Comparative Syntax* (pp. 293–372). New York NY: OUP.

Di Sciullo, A.-M. & Williams, E. (1987). *On the Definition of Word*. Cambridge MA: The MIT Press.

Good, J. (2004). Tone and accent in Saramaccan: Charting a deep split in the phonology of a language. *Lingua, 114*, 575–619.

Halle, M. & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale & S. J. Keyser (Eds.), *The View from Building 20* (pp. 111–176). Cambridge MA: The MIT Press.

Henry, A. & Tangney, D. (1999). Functional Categories and Parameter Setting in the Second-Language Acquisition of Irish in Early Childhood. In M. DeGraff (Ed.), *Language Change and Creation: creolization, diachrony and development* (pp. 239–253). Cambridge MA: The MIT Press.

Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge MA: The MIT Press.

Jackendoff, R. (2002). *Foundations of Language*. Oxford: OUP.

Keenan, E. (1981). A Boolean approach to semantics. In Jan Groenendijk, T. Janssen, & M. Stokhof (Eds.), *Formal Methods in the Study of Language* (pp. 343–380). Amsterdam: Mathematisch Centrum.

Kihm, A. (2000). Are creole languages perfect languages? In J. McWhorter (Ed.), *Language Change and Language Contact in Pidgins and Creoles* (pp. 163–200). Amsterdam: John Benjamins.

Koster, J. (1975). Dutch as an SOV Language. *Linguistic Analysis, 1*, 111–136.

Lardiere, D. (1994). The Acquisition of Word-formation Rules for English Synthetic Compounding. Ph.D dissertation, Boston University.

Lardiere, D. (1995). L2 acquisition of English synthetic compounding is not constrained by level-ordering (and neither, probably, is L1). *Second Language Research, 11*, 20–56.

Lardiere, D. & Schwarz, B. (1997). Feature-marking in the L2 development of deverbal compounds. *Journal of Linguistics, 33*, 327–353.

Lefebvre, C. (1998). *Creole Genesis and the Acquisition of Grammar*. Cambridge: CUP.

Lightfoot, D. (1999). Creoles and cues. In M. DeGraff (Ed.), *Language Change and Creation: Creolization, diachrony and development* (pp. 431–452). Cambridge MA: The MIT Press.

Migge, B. (2003). *Creole Formation and Language Contact*. Amsterdam: John Benjamins.

Mufwene, S. (2001). *The Ecology of Language Evolution*. Chicago IL: Chicago University Press.

Muysken, P. (2001). The creole languages of the Caribbean. In J. Arnold (Ed.), *A History of Literature in the Caribbean,* Volume 2: *English- and Dutch-speaking regions* (pp. 399–414). Amsterdam: John Benjamins.

Noyer, R. (1993). Optimal Words: Towards a declarative theory of word formation. Ms., Princeton, University.

Plag, I. (2003). *Word-Formation in English*. Cambridge: CUP.

Randall, J. H. (1982). Morphological Structure and Language Acquisition. Ph.D dissertation, University of Massachusetts, Amherst.

Rizzi, L. (1999). Broadening the empirical basis of Universal Grammar models: A commentary. In Michel DeGraff (Ed.), *Language Change and Creation: creolization, diachrony and development* (pp. 453–472). Cambridge MA: The MIT Press.

Roberts, I. (1999). Verb movement and markedness. In M. DeGraff (Ed.), *Language Change and Creation: Creolization, diachrony and development* (pp. 287–327). Cambridge MA: The MIT Press.

Roeper, T. (1997). Finding fundamental operations in language acquisition: Formal features as triggers. Ms. University of Massachusetts, Amherst.

Sadock, G. (1991). *Autolexical Syntax*. Chicago IL: University of Chicago Press.

Schumann, C. (1783). *Neger-Englisches Wörterbuch*. In A. Kramp (1983), *Early Creole Lexiography: A study of C.L. Schumann's manuscript dictionary of Sranan* (pp. 44–305). Ph.D Dissertation, Rijksuniversiteit Leiden.

Senghas, A., Kita, S., & Özyürek, A. (2004). Children creating core properties of language: Evidence from an emerging sign language in Nicaragua. *Science, 305*, 1779–1782.

Slobin, D. (1977). Language Change in childhood and history. In J. Macnamara (Ed.), *Language Learning and Thought* (pp. 185–214). New York NY: Academic Press.

Slobin, D. (2002). Language evolution, acquisition, diachrony: Probing the parallels. In T. Givón & B. Malle (Eds.), *The Evolution of Language out of Pre-language* (pp. 375–392). Amsterdam: John Benjamins.

Smith, N. (2001). Voodoo Chile: Differential substrate effects in Saramaccan and Haitian. In N. Smith & T. Veenstra (Eds.), *Creolization and Contact* (pp. 43–80). Amsterdam: John Benjamins

Smith, N. & Veenstra, T. (1990). *-MA nominalizations in Saramaccan and the morphology-syntax interface.* Paper presented at the ATW Morfologendag, Utrecht, May 18, 1990.

Smith, N. & Veenstra, T. (1994). Affixation in a radical creole. Paper presented at the SPCL at LSA, Boston.

Smith, N. & Veenstra, T. (1998). *Synthetic compounds in a radical creole: Abrupt versus gradual change.* Paper presented at the University of Regensburg Creole Conference, June 24–27, 1998.

Spencer, A. (1991). *Morphological Theory*. London: Blackwell.

Sportiche, D. (1999). *Reconstruction, Constituency and Morphology*. Paper presented at GLOW Berlin.

Sproat, R. (1985). On Deriving the Lexicon. Ph.D dissertation, MIT.

Tomasello, M. (2002). The emergence of grammar in early child language. In T. Givón & B. Malle (Eds.), *The Evolution of Language out of Pre-language* (pp. 307–326). Amsterdam: John Benjamins.

van den Berg, M. (2003). Early 18th century Sranan '*-man*. In I. Plag (Ed.), *The Phonology and Morphology of Creole Languages* (pp. 231–251). Niemeyer: Tübingen.

Van Dyk, P. (n.d., c 1765). Nieuwe en nooit bevoorens geziene onderwijzinge in het Baster Engels, of Neeger Engels. In J. Arends & M. Perl (1995), *Early Surinam Creole Texts. A collection of 18-thcentury Sranan and Saramaccan texts* (pp. 93–239). Frankfurt: Vervuert.

Veenstra, T. (2006). Head ordering in synthetic compounds: acquisition processes and creole genesis. In P. Bhatt & I. Plag (Eds.) *The Structure of Creole Words: Segmental, syllabic and morphological aspects*. Tübingen: Niemeyer.

Veenstra, T. (forthcoming). Creole Genesis: The impact of the LBH. In S. Kouwenberg & J. Singler (Eds.) *Handbook of Pidgin and Creole Languages*. London: Blackwell.

Verrips, M. (1998). Amsterdam-Leiden, Leiden-Amsterdam, oftewel: Waar zit je met je hoofd? In A. Bruyn & J. Arends (Eds.), *Mengelwerk voor Muysken* (pp. 12–19). Amsterdam: Publikaties van het Instituut voor Algemene Taalwetenschap, Universiteit van Amsterdam.

Wullschlägel, H. (1856). *Deutsch-Negerenglisches Wörterbuch*. Löbua: n.p. (Unchanged reprint 1965. Amsterdam: Emmering)

Zwart, J. W. (1993). Dutch Syntax. Ph.D Dissertation, Rijksuniversiteit Groningen.

The restructuring of tense/aspect systems in creole formation

Donald Winford
The Ohio State University, USA

This paper attempts to reconcile the so-called 'superstratist' and 'substratist' views on creole formation, with special attention to the emergence of tense/aspect systems in Haitian French Creole and Sranan Tongo. Creole formation involves a process of restructuring by which interlanguage grammars are created and elaborated in the course of second language acquisition. This process involves three major components: input (intake) from the TL, L1 influence, and internal developments. Differences in the degree of TL and L1 influence result in differences between Haitian Creole and Sranan Tongo in their inventory of tense/aspect categories, and the forms that express them. At the same time, both creoles display similarities due to shared substrate input as well as similar processes by which superstrate lexical forms are reanalyzed as tense/aspect markers.

1. Introduction

The nature of creole formation continues to be the subject of sometimes bitter controversy in the field of creole studies. In particular, there has been disagreement about whether creole formation is the outcome of first or second language acquisition – an issue explored in Andersen (1983) and more recently in DeGraff (1999a). Some creolists still adhere to Bickerton's Language Bioprogram Hypothesis or some version of it that ascribes the primary role in creole creation to children who appeal to innate universal principles to compensate for deficient (pidgin) input to their L1 acquisition process (Bickerton 1984, 1999). Since this theory has been sufficiently and convincingly argued against elsewhere (Arends 1995a; Roberts 2000; Singler 1992), I will not consider it here. Most creolists in fact maintain that creole formation was a type of second language acquisition, though it is clear that there are some significant differences between the two due to disparities in the nature of the two types of learning situation.

Despite this broad consensus, there is continuing disagreement among creolists over several issues. One is the question of the relative roles of superstrate and substrate input in creole formation. Second, there is the problem of what constituted the target

for creole formation. Closely related to this is the issue of whether creoles first began as close L2 approximations to the superstrate varieties, or began as community-wide pidgins that were subsequently elaborated. Creolists today are split into two broad camps with respect to these issues.

The so-called 'superstratist' view maintains that creoles began as second language varieties of the lexifier or 'superstrate' languages and gradually diverged more and more from the latter via a process of 'basilectalization' (Mufwene 1996a/b). In this view, then, most of creole grammar can be traced to the lexifier language, which was available as a target in the earlier stages of contact (Chaudenson 1992, 2001). Superstratists allow for a certain degree of substrate influence (Mufwene 1990), but appear to assign it a secondary role in creole development (DeGraff 2002).

On the other hand, adherents of the 'substratist' position claim that the major influence on the grammar of 'radical' creoles in fact came from the substrate languages (Lefebvre & Lumsden 1994; Lefebvre 1996; Lumsden 1999, etc.). The strong version of the substratist position finds expression in the so-called relexification hypothesis, which maintains that the creators of creoles were adult speakers of West African languages who attempted unsuccessfully to acquire the European target languages (TLs) to which they had highly restricted access (Lefebvre 1998: 36). Under these conditions, they resorted to a "process of vocabulary substitution in which the only information adopted from the TL in the lexical entry is the phonological information" (Lefebvre 1998: 9).

The two conflicting positions just outlined have been argued for in relation to the very same creole. Thus, writing about the TMA system of Haitian Creole (HC), DeGraff (2005: 320) argues that "all of the preverbal TMA morphemes in HC . . . can be straightforwardly traced back to seventeenth- and eighteenth-century French", and "can be analyzed as the product of, inter alia, run of the mill reanalysis or grammaticalization".

Lefebvre (1998: 111), on the other hand, argues that "the general features of the Haitian TMA system pattern on the model of Fongbe rather than French", and that

> While most of the semantic and syntactic properties of the lexical items involved in the TMA system of Haitian are derived from the corresponding lexical entries in the substratum languages, the phonological representations of these markers appear to be derived from the phonetic representations of French periphrastic and thus, lexical, expressions.

Yet another view of the emergence of creole TMA systems has been expressed by McWhorter (1999: 8), who assigns the primary role to internal developments rather than to substrate influence in the creation of the TMA system of Sranan Tongo. He argues that

> . . . there is no West African language or even class of languages, whose TMA system Sranan's could even be seen as a 'reduction' of. The Sranan TMA system dimly reflects W. African patterns but has largely developed according to its own dictates.

The present paper attempts to reconcile these conflicting positions by examining more closely the developmental stages and processes of restructuring that apply to creole formation and how they differ from (other) cases of second language acquisition (SLA). Before we proceed further, it is necessary to clarify precisely what we mean by 'restructuring'. Creolists have tended to use the term to refer to the gradual modification of earlier superstrate models, that is, the L1 varieties of the lexifier languages that were introduced into the colonies by early settlers. This implies that creole creators began with the lexifier, modifying it over time. It has also been suggested that the L2 varieties acquired by early African arrivals, and the L2 varieties of those varieties acquired by successive new arrivals, represent a continuing process of 'basilectalization' of the original superstrate target language. From this macro-level perspective, the eventual result (the creole) is seen as a 'restructured' version of the superstrate. However, terms like 'basilectalization' may be more appropriately used to refer to the gradual changes we observe in the community language over time, rather than to the psycholinguistic processes that initiate such changes on the level of individual grammars.

My own approach is concerned with the processes themselves. The restructuring I have in mind involves the ways in which individual interlanguage (IL) grammars are created and elaborated in the course of acquisition. This is the sense in which researchers in the fields of first and second language acquisition have always used the term. With respect to first language acquisition, van Buren (1996: 190) defines it as "discarding old grammars for new ones". He adds: "As soon as new relevant data are encountered, the current grammar is restructured to accommodate the new input" (ibid.). Referring to SLA, Lalleman (1996: 31) defines restructuring as "the process of imposing organization and structure upon the information that has been acquired" as new input is encountered. Hulstijn (1990: 32) defines it as "the establishment of new procedures which reorganize a body of facts and rules previously acquired". In short, restructuring has to do with the ways in which IL grammatical systems are successively expanded during the course of acquisition.

There are some basic similarities in the restructuring processes observed in creole formation and second language acquisition. First, both involve an initial or early stage of learning, in which individual learners create a highly simplified interlanguage (IL) system. This is followed by elaborative stages in which the basic IL system is expanded, drawing on three major sources of input. These include input (intake) from native and non-native varieties of the lexifier language, L1 influence, and internal developments peculiar to the IL system itself. However, significant differences in the nature of the input and the degree of access to it, among other things, lead to pronounced differences between creole formation and other types of second language acquisition. Such differences manifest themselves particularly in the greater extent and perseverance of L1 influence and internal developments characteristic of creole formation.

Tense/aspect systems reveal quite clearly the different degrees of interaction among the three factors that guide the restructuring process in creole formation, resulting in quite diverse outcomes. A closer look at the emergence of this area of creole grammar can also highlight the ways in which the process of restructuring both resembles

and differs from the elaboration of tense/aspect in more usual cases of SLA (on TMA marking see also Migge this volume).

1.1 The superstrate input to creole formation

For a long time, particularly among students of English-lexicon creoles, the conventional wisdom has been that creoles are elaborations of pidgins – the so-called 'two-stage' view of creole formation. According to this view, the superstrate input to creole formation consisted of reduced pidgin-like varieties, lacking essential components of grammar (Bickerton 1981). This conception of creole formation is expressed in Hymes' (1971: 84) characterization of creolization as "that complex process of sociolinguistic change comprising expansion in inner form, with convergence, in the context of extension in use". A radically different view has been promoted by Chaudenson (1992, 2001) and others who argue that provincial French dialects and/or relatively close second language approximations of these were the starting point of creole formation, at least in the French colonies. The two perspectives on creole formation are summed up by Alleyne (2000), who argues that two 'opposite processes' were involved in in the formation of different creoles. He notes (2000: 128):

> Whereas the kind of maximum restructuring of English represented by Saramaccan is the beginning of a historical process in the case of Saramaccan, the maximum restructuring of French represented by Haitian is the end of that historical process.

It's not clear to me what these opposing 'historical processes' refer to. I would suggest that this is a false dichotomy, which obscures significant similarities in the emergence of creoles of different lexical affiliation. It is erroneous to think of the restructuring processes that led to the genesis of HC and Saramaccan as mirror-images of each other, as Alleyne seems to suggest. The only real difference seems to lie in the kinds of superstrate input involved in each case. Whether we refer to 'basilectalization' of superstrate dialects or 'elaboration' of pidgin input, the processes of restructuring involved in the reanalysis of superstrate forms as TMA markers, remain the same.

From the point of view of the restructuring process, both scenarios for creole formation would involve an initial stage in which learners constructed a basic and highly reduced variety of the target language that has pidgin-like characteristics. This would be true whether the input to the first stages of creole formation consisted of close L2 approximations to the superstrate language, or a simplified or pidginized variety of that language. But if such different inputs were in fact the initial stage for different creoles and continued to be available as models, then the outcomes of the restructuring process would differ significantly by virtue of that fact alone. And this would explain the stark differences we find among creoles with different inputs. This seems to be the case with Haitian Creole and Sranan, as we will see.

In typical SLA, learners progress beyond their early IL system by adding more morphological apparatus, grammatical rules, vocabulary etc., the major source of

which is the TL. The degree of access learners have to the TL is in inverse proportion to the extent to which they appeal to their L1 and to creative innovation in the expansion of their IL system. The same seems to have applied in certain cases of creole formation where first and/or second language varieties of the superstrate became consolidated among a significant portion of the population, and continued to be available as targets of acquisition. In such cases, the resulting creoles were closely akin to dialects of the superstrates, as has been described for Barbadian Creole (Bajan) (Winford 2000b) and Reunion Creole (Corne 1982).

But this was not the case in other instances of creole formation. For example, with regard to Hawai'i Creole English (HCE), there appears to be agreement that the primary input came from Hawai'i Pidgin English, many of the characteristics of which persist in the creole (Roberts 1998, 2000). In this case, expansion of pidgin into creole involved use of L1 strategies by speakers of Chinese, Portuguese and other languages who were learning and using the pidgin as their primary vernacular. Hence there is ample evidence of substrate influence from these languages on HCE (Siegel 2000). Similarly, there is agreement that the main input to the formation of the Melanesian creoles was an extended pidgin that evolved out of an earlier South Seas Pidgin (Clark 1983; Keesing 1988). Siegel (1999) has demonstrated the role of substrate languages in shaping the grammar of these languages. I'll argue that the Sranan case is similar to these with regard to both the nature of the lexifier language input and the lack of full access to native varieties of the superstrate.

Another important distinction between creole formation and more typical SLA has to do with the changing nature of the input over time in the former case. As Arends (1995a), Baker (1990), Singler (1990) and others have argued, it would seem that most workers or slaves who were transported to various colonies, especially at the height of the plantation system, were attempting to learn an already established contact variety quite distinct from the lexifier languages. Indeed, in many if not most cases of creole formation, the nature and types of superstrate input changed over time, as successive waves of new learners created their own L2 versions of existing targets. In such cases, if we were to freeze the contact situation at different points in time, we would find quite different scenarios, with different targets, and hence differences in the superstrate-derived input. This presumably is what led Baker (1990) to argue that, in the formation of many creoles and 'expanded' pidgins, the true target was not the superstrate, but the emergent contact variety itself. In the light of all of this, let us now examine the emergence of the tense/aspect systems of HC and Sranan Tongo (SN).

1.2 Acquisition of tense/aspect

To place our discussion of the creole tense/aspect system in some perspective, let us first consider, very briefly, the general pattern of tense/aspect acquisition in typical SLA. Studies of SLA involving a variety of learners of different L1's attempting to acquire different L2's have demonstrated that the acquisition of L2 tense/aspect systems

follows a very similar pattern of development in all cases. According to Bardovi-Harlig (2000: 25 ff.), the following stages have been observed:

> Stage 1: The pragmatic stage. This is characterized by use of bare verbs, reliance on chronological order, and the strategy of 'scaffolding' or reliance on the other interlocutor's utterances.
>
> Stage 2: The lexical stage. In this stage, the use of bare verbs continues, and there is strong reliance on temporal and locative adverbs to convey time reference. Other strategies include the use of connectives (e.g, 'and, then'), the use of dates or days of the week, and the use of temporal verbs like 'start' and 'finish'.
>
> Stage 3: The morphological stage. Again, use of bare verbs continues, but then verbal morphology begins to appear, usually in a fixed order, depending on the target involved. For instance, the (perfective) Past tense emerges first in all cases, followed by the Imperfective Past in L2 varieties of Romance languages and the Perfect in L2 varieties of Germanic languages. [This general pattern of acquisition is based on studies by Klein (1993, 1995), and summarized in Bardovi-Harlig (2000: 119).]

This order of acquisition seems typical of learners who have both access to, and frequent opportunity to use, the target language. Clearly, it differs significantly from that found in creole formation, particularly at the morphological stage. As Bickerton (1988: 278) noted, the elimination of inflectional morphology in the early stages of creole formation results in, among other things, a loss of TMA markers. Hence these have to be reconstituted in the elaboration of creole grammar. Let us now consider the specific sources of these TMA markers, and the nature of the processes involved in their emergence in HC and SN.

2. The emergence of Haitian Creole

The French presence in Haiti dates from 1630, when freebooters established residence on Ile de la Tortue, an island off the north coast. But it wasn't until 1665 that the colony, then known as Saint Domingue, received its first French governor. The population then consisted of about 400 French planters, some slaves, plus the freebooters. By 1681, the population consisted of 4,336 whites and 2,312 slaves (Patterson 1982: 481). Soon after that, sugar plantations were established on the island, leading to rapid and massive growth of the African population until 1791, when the rebellion ended the slave trade. It seems likely that Haitian Creole emerged as a distinct language during the period 1680 to 1739 or so, and was well established by 1750 (Baker 1983: 134). As Singler (1993: 243) notes, the available records indicate that the majority of Africans brought to Saint Domingue between the 1650's and 1710 were speakers of Kwa languages, especially Ewe-Fon. The dominance of these speakers continued and perhaps grew stronger during the period from 1710 to 1739 (Singler 1993: 245). Scholars gener-

ally accept that the primary substrate input to HC formation came from Gbe languages such as Ewe and Fon.

The details of the demographic and social variables that played a role in HC genesis have been well documented elsewhere (e.g., Singler 1993, 1995, 1996), so they will not be repeated here. Two of these are especially important to our concerns here. First, the earlier period of settlement, up to 1680 or so, was characterized by small-scale farming in which there was close contact between Africans and speakers of French regional dialects. This gave many Africans, as well as children of mixed race, the opportunity to learn these dialects. Second, there continued to be a significant and increasing number of French speakers, both white and of mixed race, throughout the colonial period. Such factors help to account for the fact that early Haitian Creole was in many ways closer to the superstrate dialects than the modern language is. The creole diverged significantly from its earlier form in the course of the early plantation period, when changes were introduced primarily by Africans acquiring the creole as a second language.

2.1 The emergence of the Haitian Creole TMA system

The evidence from Haitian Creole suggests that many of its features are modeled on regional French dialects, though various kinds of simplification and reanalysis have occurred. At the same time, the gradual loss of access to such regional dialects, and the continuing process of SLA by succeeding generations of Africans in Haiti, created the conditions for significant substratum influence to affect the evolution of HC.

The major functional categories of the Haitian Creole TMA system are shown in Table 1, which is based on DeGraff (2005: 320–323) and Spears (1990). I have amended their labels somewhat.

DeGraff (2005) offers various comparisons of HC and (earlier) regional French verb structures, which demonstrate close correspondences between HC TMA markers and elements used in periphrastic strategies for marking TMA meanings in the French dialects. A few examples of this will suffice. Note first of all the clear similarities between the HC past marker *te* and French *était* 'was' in sentences like the following (from DeGraff 2005: 320).

Table 1. Haitian Creole TMA categories

Tense/aspect	
Perfective aspect	Unmarked
(Relative) Past	*te*
Prospective Future	(*a*)*pral*(*e*)
Progressive/Immediate Future	*ap*
Completive (Perfect)	*fin*(*i*)
Modal categories	
Possible Future	*va* (*a/av/va*)
Expectation/likelihood	*pu*

(1) a. HC *Li te (deja) ale.*
3SG PAST (already) go
'He had (already) gone'
b. FR *Il était (déjà) allée.*
3SG MASC was (already) go (PP)
'He had (already) gone'

A similar correspondence is found between *te* and the French past participle *été* as seen in the following (DeGraff 2005: 321).

(2) a. HC *Li te malad.*
3SG PAST sick
'S/he was/has been sick'
b. FR *Il a été malade.*
He has been sick
'He has been sick'

Detgers (2000: 150), following Chaudenson (1981: 206f.), suggests that *était* in French periphrastic constructions such as *il était à écrire* 'he was writing' was the source of Past *te* in French creoles. It seems clear that *te* has its source in the French past imperfect *était/étais*, with possible reinforcement from past participial *été.*

Similar correspondences can be found between HC modal *pu* and its French cognate, the preposition *pour* 'for' and between Future *va* and French *va*(*s*), the present singular forms of *aller* 'to go' used in the Future construction *aller* + V 'be going to V.' The following examples from DeGraff (2005: 321) illustrate this.

(3) a. *Mwen pou marye semen pwochèn.* (HC)
1SG for marry week next
b. *Je suis pour me marier la semaine prochaine.* (Canadian French)
1SG am for me marry the week next
'I am to get married next week'

(4) a. *Ou (a)va ale demen.* (HC)
You FUT go tomorrow
'You will go tomorrow.'
b. *Tu vas aller demain.* (French)
'You will go tomorrow.'

Similar (regional) French cognates can be found for other HC TMA markers. For example, Progressive marker *ap*(*e*) has its source in the preposition *après*, employed in the earlier French construction *être après à* +V 'to be V-ing.' HC Prospective (*a*)*pral*(*e*) can be traced to the progressive construction *après* (*de/à*) *aller* + V 'to be going to V'. terminative perfect *fin*(*i*) similarly derives from the lexical verb *finir*.

Table 2 summarizes the correspondences between the TMA markers of HC and their regional French cognates.

Table 2. Sources of main HC TMA markers

HC catetory	HC marker	Regional French sources
Perfective	Unmarked	Infinitival/3rd sing/particle
(Relative) Past	*te*	Imperf. *était* / PP été
Prospective Future	(*a*)*pral*(*e*)	*après* (*de/à*) *aller*
Progressive/Immediate Future	*ap*	*être après à* + V
Completive (Perfect)	*fin*(*i*)	*finir* 'finish'
Possible Future	*va* (*a/av/va*)	*va*(*s*) + V
Expectation	*pu*	*être pour* + V

While researchers agree that the phonetic shapes of these TMA markers derive from French cognates, there is strong disagreement concerning the relative roles of superstrate and substrate input in shaping the semantics and syntax of the markers. Let us therefore consider each of these influences, as well as the contribution of internal developments in the emergence of HC tense/aspect.

2.2 The superstrate input to HC

Chaudenson (1995), Fattier (1998), DeGraff (2005), and others have argued that most of the HC TMA markers derive their semantic and distributional properties from those of their French cognates. This would apply especially to markers such as past *te*, future *a/av/va*, prospective (*a*)*prale* and modal *pou*. The correspondences outlined earlier suggest that this claim is at least partly true, which is in keeping with the fact that the superstrate input to HC came from first and second language varieties of regional French that remained available as models, at least during the first stages of HC formation. But the French input alone cannot explain various characteristics of the HC tense/aspect system, which have to do with internal developments and substrate influence.

2.3 Internal developments

The internal developments we are concerned with here are those that contributed to the earlier stages of HC creation, up to the point in the early 18th century when the language crystallized as a medium of communication quite distinct from its French superstrate. All of the markers we have discussed so far were established by this time, though it appears that they did not all emerge simultaneously (Baker 1995). Some of the developments that characterized these early stages included processes of simplification and reduction leading to the loss of inflectional and other non-salient elements of the French verb complex. Such processes are, of course, typical of both first and second language acquisition, and of contact-induced change; they constitute what Chaudenson (1992: 152) refers to as 'natural developments'.

In addition to such developments, HC formation also involved a learning strategy in which lexical content items such as *finir* and *après*, and other salient, quasi-

grammatical forms such as *était/été* and *pour*, were chosen to express grammatical notions associated with tense, mood, and aspect categories. DeGraff (2005: 320), Mufwene (2001: 54) and others refer to this process as 'grammaticalization'. Detgers (2000: 145) questions the general application of this term to the emergence of HC TMA markers, arguing that only some of them were due to grammaticalization in the strict sense (i.e., the reinterpretation of lexical items as grammatical elements), while others were due to 'reanalysis'. An example of the former is the perfect marker *fin*(*i*), whose French cognate, *finir* 'to finish,' had no grammatical function. Cases of reanalysis, by contrast, involve markers that are etymological continuations of forms that already had grammatical or quasi-grammatical functions in the superstrate. Examples include future *a/va/ava*, past *te*, and prospective (*a*)*pral*(*e*). Modal *pu*, on the other hand, seems to have involved elements of both processes.

Detger's distinction is useful because it distinguishes continuities from the superstrate that are due to reanalysis from cases of grammaticalization, which tend to proceed over time under internal motivation. But it is clear that some cases of the latter process, such as perfect *fin*(*i*), occurred early in creole formation, and were very likely accelerated due to substrate influence (Bruyn 1996). It should however be noted that several later developments in the verb complexes of HC and other creoles are due to gradual grammaticalization of the more usual type, as Alleyne (2000) has pointed out. This reinforces the need for caution in using the contemporary structure of creole TMA systems as the basis for discussion of their genesis.

A final example of internal developments in HC is the gradual emergence of combinations of markers to express more complex tense/aspect notions. As Baker (1995) has shown, most of the combinations attested in earlier HC texts are found much later than the single markers. They include combinations like past + future, past + future + progressive, etc, which might have been due, at least in part, to substrate influence from Gbe, which employs similar, though not identical, combinations. Such sequences are also found in Sranan, and I will reserve further discussion of them for later.

It is clear that the processes of change by which French items were reinterpreted as TMA items were peculiar to the sociohistorical circumstances in which HC and other French-lexicon creoles emerged. Such developments are in stark contrast with the usual pattern of tense/aspect acquisition in typical L2 French. Schlyter (1990) provides the following picture of this pattern:

Pattern of acquisition of tense/aspect in L2 French.
Stage 1: One or two basic forms of verbs with variable use.
Stage 2: The *passé composé* emerges, though not entirely productive.
Stage 3: Use of *veux* + infinitive or *va* + infinitive to express future meaning.
Stage 4: Clear cases of the imparfait emerge.
Stage 5: The pluperfect, conditional and subjunctive categories emerge.

Stages 1–3 (corresponding roughly to the lexical stage) bear much resemblance to what we would expect in the early stages of acquisition in the Haitian context. But traditional SLA parts company with creole formation in the morphological stages, which corre-

spond roughly to the stage of restructuring. The categories in stages 4 and 5 above have no morphological expression in HC, by contrast with advanced L2 French. Moreover, as Mather (1995) and others have noted, L2 varieties of French (and other European languages) manifest few, if any, instances of preverbal marking, of the sort found in HC. This can be explained, in part, as the result of Gbe influence on HC. As Mather (2000: 259) points out:

> Once the French periphrastic constructions were stripped of their inflectional endings by the first generation of creole speakers, they could be reinterpreted as preverbal TMA markers by adult and children speakers of Kwa languages, who identified them with their own L1 TMA markers.

By contrast, the Klein/Purdue studies of SLA in Europe found a striking lack of L1 influence on L2 TMA where one might expect it (see Klein 1993, 1995; and Klein & Purdue 1997 for more details). Thus, Klein et al. (1995: 278) conclude that there is no significant L1 influence on the acquisition of temporality. The strong L1 influence on HC tense/aspect is a result mainly of the continuing acquisition of the then available contact variety by successive groups of newly arrived Africans.

2.4 Substrate influence on HC

We saw earlier that Kwa, and in particular, the Gbe sub-family, were the primary substrate input to HC formation. Almost all of the controversy surrounding the emergence of the tense/aspect system of this creole revolves around the nature and extent of that substrate influence. So-called superstratists acknowledge very little input from African learners' L1's. DeGraff (2005) allows that substrate influence did play some role in two aspects of the HC tense/aspect system. On the one hand, the SLA-related strategy of employing infinitival-like verb forms would have been encouraged by the fact that Kwa does not employ tense and agreement affixes (DeGraff 2005: 327). On the other hand, the preference for preverbal marking based on French periphrastic constructions would have been favored by the fact that Gbe also has non-affixed preverbal TMA markers.

These acknowledgements of (limited) Gbe influence are in stark contrast to the view of Lefebvre (1996, 1998), who claims that most of the semantic and syntactic properties of HC TMA markers derive from those of corresponding substrate (Fongbe) markers. Space does not permit a critique of Lefebvre's position here. It has been argued elsewhere, e.g., by Winford (2002) and Migge and Winford (2003) that many of her claims concerning (especially) semantic correspondences between Fongbe and HC markers are questionable. Moreover, as DeGraff (2005: 328) and McWhorter (1999) have both noted, there are also significant differences between Gbe and HC in the inventory and configuration of TMA markers, while "Fongbe verbs do manifest inflectional and syntactic processes that are *not* attested in HC" (DeGraff 2005: 305). Hence the HC TMA system can hardly be regarded as a replica of the Gbe system. Further research is needed to clarify the issue.

The general conclusion to be drawn from our overview of HC tense/aspect is that both the superstratist and substratist accounts of its emergence have merit. Clearly some compromise between the two positions is necessary to account for the developments we have discussed.

We would expect that, in cases where superstrate input is even more limited, creole creators would compensate for this by drawing more heavily on L1 knowledge as well as the internal resources of their developing IL system. A case in point is Sranan Tongo.

3. The emergence of tense/aspect in Sranan Tongo

Like Haitian, Sranan Tongo employs preverbal free forms to express temporal, aspectual and modal meanings. One exception is the completive marker *kaba*, which always occurs in VP-final position. The inventory of the major tense/aspect categories and the forms that express them in SN are shown in Table 3. Note that potential *sa* is more of a modal than a marker of just future time reference in contemporary Sranan, but appears to have had future marking as its primary function in earlier SN.

Sentences (5–10) illustrate the use of each of the respective tense-aspect categories. The relevant forms are in boldface.

(5) *A* ***djuku*** *wan man* ***boro*** *en here bere.*
3SG. stab Art. man cut.open 3SGPOSS whole belly.
'He stabbed a man and cut open his entire belly'

(6) *Wan tu fu den pikin fu owma* ***e*** *wroko gron now ooktu?*
One two of the-PL child of granny IMP work ground now too
'Are some of granny's children also cultivating the land now too?'

(7) *A alen disi kan stop now. Yongu, a kon tumsi furu* ***kaba****, yere.*
The rain this can stop now. Man it come too full already, hear
'This rain can stop now. Man, it has already rained more than enough'

(8) *A* ***ben*** *taigi mi a o kon na fesisey baka. Mi no sabi*
he PAST tell me he FUT come LOC front.side back. I NEG know
efu a go ete.
if he go yet
'He told me he would come to the front again. I don't know if he's gone yet'

Table 3. Major Tense/Aspect categories in Sranan (Winford 2000a)

Aspect	Perfective	ø (the unmarked verb).
	Imperfective	*e*
	Completive Perfect	VP-final *kaba*.
Tense	Relative Past.	*ben*
	Predictive Future	*o*
	Potential Future (modal)	*sa*

Table 4. Tense/Aspect categories in Sranan and their sources

Sranan category	Marker	Source
Perfective	Unmarked	Bare verb
Imperfective	*e* < *de*	English *there*
Completive Perfect	*kaba* (VP-final)	Portuguese *acabar* 'finish'
Relative Past	*ben*	Eng. *been*
Predictive Future	*o*	Eng. *go*
Potential Future	*sa*	Dutch *zal*

(9) *Efu yu no wroko, dan yu 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?'

(10) *Dan te mi miti en mi* ***sa*** *aksi en.*
Then when I meet him I POT ask him
'Then when I meet him I will ask him'

It is clear from the above examples that, unlike the preverbal markers of Haitian Creole, those in Sranan Tongo have no cognates in any English tense/aspect markers. One possible exception to this is the potential marker *sa*, which some have claimed to be a form of English *shall*. However, it is much more probable that it derives from Dutch *zal*. In fact, in early SN texts such as Van Dyk (1765), the future marker is written variously as *zal, sal, za*, etc. Table 4 provides an overview of the actual sources of the Sranan tense/aspect markers.

This presents a very different picture from that we saw earlier for Haitian creole. In the first place, there are few, if any distributional or semantic similarities between the SN markers and their English cognates. Second, two markers, *sa* and *kaba*, have been adopted, not from English, but from Dutch and Portuguese respectively. This can only be explained in terms of the limited access to and input from, varieties of English among the Africans who created Sranan Tongo. Needless to say, the Sranan situation is in stark contrast with more typical cases of L2 acquisition of English tense/aspect. An illustration of the latter is shown in Table 5.

We can now consider the circumstances that contributed to the radical divergence from superstrate dialects that is manifested in Sranan tense/aspect.

3.1 The superstrate input to Sranan formation

Since the precise details of the English input to Sranan are not well known, some consideration of the historical background to this creole's genesis is in order. Sranan in fact shares much with Haitian Creole with regard to the circumstances of its origin, but there are some significant differences between the two in the demographics and nature of the contact between Europeans and Africans. Such differences may help to explain the different degrees to which each diverged from its lexifier language.

Table 5. Order of acquisition of English verbal morphology
Based on Klein's (1993, 1995) study of Lavinia (L1 = Italian)

Period in UK	Features acquired
6 months.	Emergent use of 3pers. *-s* and present copula.
7 months.	3 irregular pasts (*said, went, was*).
8 months.	4 tokens of Present Perfect (no contrast with Past).
	1 token of Future. Increased use of V-*ing*.
11 months.	Past expressed mostly by irregular pasts.
13 months.	First use of regular past.
16 months.	Increased use of regular past.
	Use of Past Progressive and Present Perfect.
17 months.	Several correct uses of Progressive.
21 months.	First clear use of Pluperfect.

3.2 Historical background

The English colonized Suriname in 1651, but ceded it to the Dutch in 1667. During this initial period, English planters introduced the plantation system, bringing with them slaves most of whom had already lived in other colonies such as Barbados. By 1667, the numbers of Europeans in the colony numbered roughly 1500, while Africans numbered two to three thousand. In the 1660's, a group of roughly 200 Portuguese-speaking Sephardic Jews came to Suriname and established plantations up river from Paramaribo, the coastal capital. The Portuguese-based contact variety used on these plantations provided part of the lexical input to Saramaccan, a creole developed by maroons in the interior. After the Dutch took control, the number of English settlers declined from approximately 1500 in 1666 to only about 38 in 1680 (Voorhoeve & Lichtveld 1975: 2–3). Planters of various other nationalities came to replace them, including persons of English, Dutch, German and Jewish origin. Though many English planters remained after the Dutch assumed control, most of them had left by 1695, taking with them only those slaves acquired before the colony was ceded to the Dutch. According to Postma (1990: 185), there were roughly 379 Europeans and 4,618 Africans in the colony in 1695.

Since the earliest Sranan texts date back to the 1710's (Van den Berg 2000), we can assume that the creole emerged in its first form sometime during the period 1651 to 1700, and was most probably well established by the time most of the English left. The early superstrate input to Sranan must have come from southern and southwestern English dialects spoken by planters and indentured servants, as well as the second language or pidgin varieties of these spoken by the slaves they had brought with them. After the English left, Dutch planters became even more involved in plantation management, though their slaves spoke some version of early Sranan. Exposure to Dutch, however, must have played some role in the development of Sranan, especially after 1680.

From 1675 on, there was increasing importation of African slaves, leading to a situation where the population consisted primarily of new slaves and a much smaller minority of Europeans, including English, Dutch and Portuguese, among others. The ratio of Africans to Europeans increased from 2/3:1 in 1679 to 12:1 by 1680 and during the 1680s the ratio of old to new slaves decreased from nearly 7:1 to nearly 2:1. By 1720, according to Postma (1990: 185), there were roughly 935 Europeans and 13,604 Africans in the colony. As was the case in Haiti, the latter were clearly dominated by speakers of Gbe varieties in this period. They made up about 70% of all the slaves brought to Suriname during the early 18th century. Less than 20% were speakers of Kikongo varieties.

The Sranan situation differed in two fundamental ways from that in 17th to early 18th century Haiti. First, the plantation system was introduced very early after British colonization, so that there was no protracted period of small farming in which Africans were in close contact with Europeans. Second, perhaps the most crucial factor in the emergence of Sranan was the very early withdrawal of the vast majority of English-speaking planters and the slaves who originally came with them, within roughly thirty years of the colony's inception in 1651. This meant that the major input to new arrivals from Africa after 1680 came from pidginized or highly changed second language varieties of English (Migge 1998). It seems likely that these contact varieties did not have a developed TMA system that could serve as a model for learners. This meant that the TMA system of early Sranan had to be built up practically from scratch, via reanalysis of available English lexical items under the influence of the substrate languages. In addition, the creole adopted its future marker *sa* from Dutch and its completive perfect marker *kaba* from a Portuguese-lexicon pidgin/creole. This in part explains why the Surinamese creoles diverge so radically from their original English sources. In short, what most distinguishes Suriname from Haiti is the almost complete withdrawal of lexifier language models (including close approximations acquired by many Africans) in the former colony.

3.3 Substrate influence on Sranan tense/aspect

The role that substrate influence played in the emergence of Surinamese creole TMA has been discussed in some detail by Migge and Winford (2003). They demonstrate that tense/aspect categories like the perfective (the unmarked verb), the VP-final completive marker *kaba*, and the imperfective marker (*d*)*e* all have clear models in the TMA systems of Gbe languages. There are also strong parallels between the creoles and their Gbe substrates in the syntax of TMA, for instance in the combinatory possibilities among the markers themselves, and the ordering of the markers. There is also the obvious similarity between the two groups of languages in their use of periphrastic rather than (bound) morphological means of conveying TMA notions.

With regard to the choice of forms to express the categories, we can note the following:

- The perfective category is expressed by the unmarked verb, and represents situations viewed as unanalyzed wholes, yielding a present interpretation with statives and a past interpretation with non-statives when the reference point is speech time S.
- Imperfective (progressive and habitual) meanings are expressed by a form (*d*)*e*, which is homophonous with a locative copula.
- Completive aspect is conveyed by a form *kaba* (< *acabar* 'finish'), which is homophonous with a verb meaning 'finish.'

All of these categories have similar uses and expression in Gbe languages, suggesting that substrate influence played a role in their emergence. To demonstrate this, I present a brief overview of Gbe tense/aspect categories in Table 6. As can be seen, all of the languages share more or less the same inventory of categories, though there are some differences. For example, completive aspect is conveyed by preverbal markers in Maxi, Xwela, and Xwla (Pattern A), but by a post verbal marker in Aja, Gen, and Waci (Pattern B). Note, however, that all Gbe languages employ two structures for the expression of progressive meaning. For convenience, only general patterns are represented, and some forms used to express the categories are not all listed in cases where their phonological shapes vary significantly (for example, forms of 'be').

3.4 Comparing perfective in Gbe and Sranan

The perfective category, expressed by the unmarked verb, is used in a more or less identical range of meanings and uses in Gbe and Sranan. The following example sentences were elicited with the help of native speakers in Benin and Suriname (Migge & Winford 2003).

In all these languages, the unmarked verb conveys the sense of 'present' with statives (1) and 'simple past' with non-stative verbs (2) in the default cases, where the point of reference is S.

(11) Waci *ɖàví á lɔ̀n àtɔ́ á kù nɔ́ á.*
child DET love father DET with mother DET
'The child loves his father and his mother'

SN *A boi lobi a umapikin.*
DET boi love DET girl
'The boy loves the girl'

(12) Aja *E cúcú èyi xɔ́mɛ.*
he clean his room
'He cleaned his room'

SN *A krin en kamra.*
He clean/arrange his room

Table 6. Tense/Aspect categories in Gbe languages

Form	Category	Meanings/Uses
Tense		
lá/ná/á + V	Future	Later time reference.
Aspect		
ø	Perfective	States or events seen as unanalyzed wholes. Simple past with non-statives, present with statives (when reference point is S).
	Completive	Situations seen as completed. 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.
kò/mɔ/nɔ̀ + V (Maxi/Xwela/Xwla)	Pattern (*a*)	
VP + vɔ̀ (Aja, Gen, Waci)	Pattern (*b*)	
	Progressive	Events in progress.
'Be' VV Part	Structure (a)	In cases where the Prog. Marker immediately precedes the verb, e.g. intransitives or transitive verbs taking a pronominal object (SVO order).
'Be' XP V Part	Structure (b)	In other transitive sentences.
	Habitual	Customary or habitual situations.
V + *na/nɔ*	Pattern (a)	(Aja, Gen, Wací)
nɔ + V	Pattern (b)	(Maxi, Xwla)
High tone on V		(Xwela)
	Prospective	Events about to occur.
'Be' (XP) nà V (Fongbe, Gungbe etc.)	Pattern (a)	
'Be' (XP) V gé/gbé (Ewe)	Pattern (b)	
'Be' (XP) já V (Waci)	Pattern (c)	

The unmarked verb is also used in all the languages to convey the sense of current relevance, in much the same way as the English perfect does. This is illustrated in the following examples:

(13) Aja *Wò wu àxɔ̀su lɔ.*
they kill king DET
'They killed the king'
SN *Den kiri a kownu.*
'They killed the king'

The close similarities in meaning and use of the unmarked verb suggest that Gbe influence played a primary role in the emergence of the perfective category in the Surinamese creoles.

3.5 Comparing the completive in Gbe and Sranan

As Table 6 shows, there are variations in the expression of the completive category across the Gbe languages. Varieties such as Maxi, Fon, Xwela, and Xwla employ preverbal markers to express the sense of perfect, as well as a construction with a VP-final form derived from the verb 'finish', to express the sense of completion of the event expressed by the main verb. Western Gbe varieties like Aja, Gen and Waci employ only the latter kind of construction. This is also a pattern found in Sranan and other Surinamese creoles, though VP-final *kaba* has been grammaticalized into a more adverbial marker conveying the sense of 'already'.

(14) Aja *àxɔ̀su lɔ à, e vá ɖó vɔ̀.*
king the TOP he come arrive COMPL
'As for the king, he has already come'
SN *A kownu doro kaba.*
DET king arrive COMPL
'The king has already come'

The completive is also used more or less uniformly in interrogative sentences in which the sense of 'already' is expressed, as the following illustrate:

(15) Gen *O kpɔ́ fofó nyɛ vɔ̀ a?*
you see brother my COMPL Q
'Have you already seen my brother?'
SN *yu miti mi brada kaba?*
You meet my brother COMPL
'Have you met my brother?'

The close similarities in these cases are reinforced by the fact that the forms used to instantiate the category completive in the Surinamese Creoles and in several of the Gbe varieties (Aja, Gen, Waci) are also used as main verbs that mean 'finish'. The following examples illustrate.

(16) Gen *gàlí-á vɔ̃* (Gengbe. Jondoh 1980: 50)
gali-the finish
'There's no more gali'
SN *Moni kaba.*
money finish
'There is no more money'

The strong similarities in the use of *kaba* 'finish' as a main verb and a marker of completive aspect in the Surinamese creoles are striking, and would therefore seem to be the result of substrate influence.

3.6 The expression of 'imperfective' meaning in Gbe and Sranan

In Sranan, aspectual marker *e* expresses a range of meanings, including "durative, iterative and continuous" (Voorhoeve 1957: 376; Seuren 1981: 1052). Bickerton (1981) analyzed it as instantiating an aspectual category that he labeled 'non-punctual'. Winford (2000a) provides clear evidence that it represents an imperfective aspectual category, whose primary use is to express both 'habitual' and 'progressive' meanings, and which can express other secondary meanings as well.

The Gbe languages have no imperfective category, distinguishing between a habitual and a progressive. It is the latter category that first provided the model for SN *de*, which is used in the early SN texts as a locative copula, but also as an apparent progressive marker. The progressive in Gbe languages is generally expressed by a copula that takes what appears to be a nominalized VP followed (in most cases) by a particle of some kind.

(17) Waci *mə̀ lə̀ wòmá wlɔ̀n ŋ̀*
I be-at letter write PART
'I am writing a letter'

(18) Xwla *uòn lè nú wlàn nù.*
I be-at thing write PART
'I am writing something'

In Sranan, the imperfective marker *e* expresses progressive meaning, as the following example illustrates.

(19) SN A no e skrifi brifi, a e sribi.
he NEG PROG write letter he PROG sleep
'No, he's not writing letters, he's sleeping'

The important thing for our purposes is the fact that the first element in all these constructions (*lé, lə̀, ɖ oò,* etc. in Gbe, (*d*)*e* in Suriname) is identical to the locative copula (Jondoh 1980: 37; Lefebvre 1996: 269).

(20) Gen *é lè ekplɔ̂-a ji.* (Gengbe. Jondoh 1980: 46)
it be table-the on.
'It's on the table.'

SN A pikin de a oso
DET child COP LOC house
"The child's at home"

It appears that the emergence of *de* as a locative copula in the early plantation creole was the trigger for its extension to the marking of progressive meaning. The model for this was the fact that the locative copula of the Gbe languages had the same function. Later, this progressive marker evolved into a marker of imperfective aspect (see below). Hence this category can be traced ultimately to Gbe influence.

The potential puture *sa* almost certainly derived from Dutch *zal*, and appears to be a case of borrowing from that language. Even so, however, there are striking parallels between the use of *sa* and the use of the (Western) Gbe potential future markers *lá, ná, á*, which express the sense of future possibility or potential (Essegby to appear; Migge & Winford 2003). There is also evidence of significant Gbe influence on the semantics and uses of several modal auxiliaries, including those that convey learned ability, positive and negative physical ability, need, obligation and others (Migge 2004; Van den Berg 2004).

Finally, the ordering of preverbal auxiliaries in SN also matches that of its Gbe substrate to a significant degree. For instance, it is well known that SN, like other radical creoles, displays a tense-mood-aspect order of auxiliaries (among others), as in the following examples from my data:

(21) Wel, dan granpapa ben sa e gi den [owru tori – DW]
well, then granpa T M A give them [old story]
'Well, then grandad would have (habitually) told them'

(22) En dan a man ben musu e breiti
And then the man T M A glad
'So the man must have been happy'

Bickerton (1984) ascribed this TMA ordering to the workings of a bioprogram. But it has a more obvious model in the Gbe substrates. Jondoh (1980: 52) informs us that, in Gengbe, "the order of constituents in AUX is generally, Tense, Modality, Aspect". She provides the following examples:

(23) é la ténú nɔ̌ ɖù gàlí. (Gengbe: Jondoh, p. 52)
he FUT can PROG eat gali.
'He will be able to eat gali'

Other auxiliary orderings which are found both in SN and Gengbe include the combination of modal and progressive (Jondoh 1980: 33) and future and progressive (p. 30). In both languages too, the negative auxiliary precedes all other auxiliaries in the verb phrase (p. 65).

This is not to say that the structure of AUX in SN is identical to that of Gengbe or other Gbe dialects. There are several significant differences, but space does not permit full discussion of them here. Suffice it to say that much of the syntax of auxiliary ordering in SN can be explained in terms of influence from the Gbe substrate. It is possible that Akan and Kikongo played some role as well, since they manifest some similarities to Gbe in their auxiliary combinations. It is also quite likely that many aspects of the syntax of the verb in SN are due to innovations and internal developments in the course of the language's development. Future research will no doubt clarify this.

The facts outlined here support the claim that the overall structure of the Sranan verb complex - the preference for periphrastic expression, the patterns of ordering and

other syntactic properties of the auxiliaries – is patterned primarily after the dominant Gbe substrates.

3.7 Internal developments in Sranan tense/aspect

Internal developments also played a role in the emergence of certain TMA categories, and in the further evolution of others. For example, relative past *ben* clearly derives from grammaticalization of *been*, a process that has also been attested in various English-lexicon creoles, including most Caribbean creoles, and Pacific creoles such as Hawai'i Creole English and varieties of Melanesian Pidgin. The relative past emerged quite early in Sranan, and is attested in one of the earliest SN texts, the Herlein fragment of 1718. There is clearly no substrate influence involved here, since Gbe languages have no past category.

A similar case of grammaticalization is the emergence of predictive future *o* from *go*, another development that is widely attested in other English-lexicon creoles, and parallels the emergence of *a/va/ava* as a future marker in French creoles. There is evidence, however, that future *o* emerged much later than past *ben*, and was not part of the early (late 17th century) Sranan TMA system. In earlier 18th century texts such as Van Dyk's manual (1765), *go* is used mainly as a main verb of motion, while future meanings are expressed by *sa*. But later texts such as Schumann (1783: 50) contain some instances of the use of preverbal *go* as an apparent future auxiliary, though arguably it could still be analyzed as a verb (Van den Berg 2004). Now, of course, *o* is well established as a marker of future tense. Again, this seems to be a purely internal development in which substrate influence played no role.

Finally, the emergence of earlier progressive (*d*)*e* as an imperfective marker seems to have been due to internal processes of grammaticalization over time. This involved extension of the meaning of *de* to cover habitual and generic meanings as well. This kind of development has parallels in the development of progressives into imperfectives cross-linguistically (Bybee et al. 1994: 141).

Contemporary Sranan also has other TMA markers that evolved via gradual grammaticalization (e.g. modal *man* 'be able to'), or were borrowed from Dutch (e.g., modal *mag* 'may'). Fuller details of these and other developments in the restructuring of Sranan TMA can be found in Migge (2004) and Migge and Winford (2003).

The creation of Sranan grammar was clearly a gradual process, in which successive generations of learners contributed in different ways to the elaboration and systematization of the grammar. The evidence suggests that several TMA markers emerged after the initial stages of creole formation in the late 17th to early 18th centuries. This seems to be true also of combinations of TMA markers, few of which are attested in early texts such as van Dyk.

4. Conclusion

The foregoing comparison of the emergence and development of TMA systems in Haitian Creole and Sranan Tongo demonstrates that no single formula can be found to explain creole formation. There are some respects in which the process is similar to that found in cases of second language acquisition in 'natural' settings, but there are significant differences as well, some of which adherents of the 'superstratist' position have pointed to. For instance, there are differences in the nature of the target language and the kinds of input from that source. Another major difference lies in the perseverance of L1-based strategies and other internal innovations in creole formation, by contrast with SLA, which, as it progresses, typically involves replacement of such strategies (and other compensatory ones) by those adopted from the TL. Creoles whose creators have had more access to superstrate sources exploit those resources more fully, and as a result, approximate superstrate grammars more closely than others. Some, like Sranan Tongo, depart more radically from the lexifier language because of the need to rely more on L1 knowledge and internal innovations, due to restricted availability of native superstrate models. In general, however, it seems reasonable to claim that creole formation was essentially a process of SLA with (usually) restricted TL input under unusual social circumstances.

With regard to the processes of acquisition, the initial stages of creole formation involve processes of reduction and simplification found in early SLA. But the restructuring of the initial IL system takes a very different path in creole formation because of the nature of the (changing) input, and lack of access, especially by later African arrivals, to native varieties of the superstrate. We've seen that the restructuring of tense/aspect systems involves processes of reanalysis due to "transfer" or substrate influence, and processes of internal change, which sometimes act in concert with the former.

From this perspective, the substratist and superstratist views of creole formation both have some validity. In fact, they complement each other, and there is no need to convert them into matters of opposing dogma. The disagreement between these camps diminishes in importance once we recognize the competing and complementary roles of substrate influence, superstrate input and internal innovation in the processes of restructuring that gave rise to creoles.

References

Alleyne, M. (2000). Opposite processes in creole formation. In I. Neumann-Holzschuh & E. W. Schneider (Eds.), *Degrees of Restructuring in Creole Languages* [Creole Language Library 22] (pp. 125–133). Amsterdam: John Benjamins.

Andersen, R. W. (1983). Transfer to somewhere. In S. Gass & L. Selinker (Eds.), *Language Transfer in Language Learning* (pp. 177–201). Rowley MA: Newbury House.

Arends, J. (Ed.). (1995a). *The Early Stages of Creolization* [Creole Language Library 13]. Amsterdam: John Benjamins.

Arends, J. (1995b). Demographic factors in the formation of Sranan. In J. Arends (Ed.), *The Early Stages of Creolization* [Creole Language Library 13] (pp. 233–285). Amsterdam: John Benjamins.

Baker, P. (1983). Assessing the African contribution to French-based Creoles. In S. Mufwene (Ed.), *Africanisms in Afro-American varieties* (pp. 123–155). Athens GA: University of Georgia Press.

Baker, P. (1990). Off target. *Journal of Pidgin and Creole Languages, 5*, 107–119.

Baker, P. (1995). Some developmental inferences from the historical studies of pidgins and creoles. In J. Arends (Ed.), *The Early Stages of Creolization* [Creole Language Library 13] (pp. 1–24). Amsterdam: John Benjamins.

Bardovi-Harlig, K. (2000). *Tense and Aspect in Second Language Acquisition: Form, meaning and use*. Oxford: Blackwell.

Bickerton, D. (1981). *Roots of Language*. Ann Arbor MI: Karoma.

Bickerton, D. (1984). The language bioprogram hypothesis. *The Behavioral and Brain Sciences, 7*, 173–188.

Bickerton, D. (1988). Creole languages and the bioprogram. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey*: Vol II. *Linguistic theory: Extensions and applications* (pp. 268–284). Cambridge: CUP.

Bickerton, D. (1999). How to acquire language without positive evidence: What acquisitionists can learn from creoles. In M. DeGraff (Ed.), *Language Creation and Language Change: Creolization, diachrony and development* (pp. 49–74). Cambridge MA: The MIT Press.

Bruyn, A. (1996). On identifying instances of grammaticalization in creole languages. In P. Baker & A. Syea (Eds.), *Changing Meanings, Changing Functions: Papers relating to grammaticalization in contact languages* (pp. 29–46). London: University of Westminster Press.

Bybee, J., Revere P., & William, P. (1994). *The Evolution of Grammar: Tense, aspect and modality in the languages of the world.* Chicago IL: The University of Chicago Press.

Chaudenson, R. (1981). *Textes creoles anciens* (*La Réunion et Île Maurice*). *Comparaison et essai d'analyse.* Hamburg: Helmut Buske.

Chaudenson, R. (1992). *Des îsles des hommes, des langues*. Paris: L'Harmattan.

Chaudenson, R. (1995). *Les Créoles*. Paris: Presses Universitaires de France.

Chaudenson, R. (2001). *Creolization of Language and Culture*. (Revised version of Chaudenson 1992, in collaboration with Salikoko Mufwene.) London: Routledge.

Clark, R. (1983). Social contexts of early South Pacific pidgins. In E. Woolford & W. Washabaugh (Eds.), *The Social Contexts of Creolization* (pp. 10–27). Ann Arbor MI: Karoma.

Corne, C. (1982). A contrastive analysis of Réunion and Isle de France Creole French: Two typologically diverse languages. In P. Baker & C. Corne (Eds.), *Isle de France Creole: Affinities and origins* (pp. 7–129). Ann Arbor MI: Karoma.

DeGraff, M. (1999a). *Language Creation and Language Change: Creolization, diachrony and development*. Cambridge MA: The MIT Press.

DeGraff, M. (1999b). Creolization, language change, and language acquisition: An epilogue. In Michel DeGraff (Ed.), *Language Creation and Language Change: Creolization, diachrony and development* (pp. 473–543). Cambridge MA: The MIT Press.

DeGraff, M. (2002). Relexification: A reevaluation. *Anthropological Linguistics, 44*, 321–414.

DeGraff, M. (2005). Morphology and word order in 'creolization' and beyond. In G. Cinque & R. Kayne (Eds.), *The Oxford Handbook of Comparative Syntax* (pp. 293–372). New York: OUP.

Detgers, U. (2000). Two types of restructuring in French creoles: A cognitive approach to the genesis of tense markers. In I. Newman-Holzschuh & E. W. Schneider (Eds.), *Degrees of Restructuring in Creole Languages* [Creole Language Library 22] (pp. 135–62). Amsterdam: John Benjamins.

Essegbey, J. (to appear). On why the *a*-morpheme is not a tense marker. In F. Ameka & M. E. Kropp Dakubu (Eds.), *Aspect and modality in Kwa languages*.

Fattier, D. (1998). Contribution à l'Étude de la Genèse d'un Créole: L'Atlas Linguistique d'Haïti, Cartes et Commentaires. Ph.D dissertation, Université de Provence. (Distributed by Presses Universitaires du Septentrion, Villeneuve d'Ascq, France.)

Hulstijn, J. (1990). A comparison between information processing and the analysis/control approaches to language learning. *Applied Linguistics, 11*, 30–45.

Hymes, D. (1971). General conceptions of process: Introduction. In D. Hymes (Ed.), *Pidginization and Creolization of Languages* (pp. 65–90). Cambridge: CUP.

Jondoh, E. E. A. (1980). Some Aspects of the Predicate Phrase in Gengbe. Ph.D dissertation, Indiana University.

Keesing, R. (1988). *Melanesian Pidgin and the Oceanic Substrate*. Stanford CA: Stanford University Press.

Klein, W. (1993). The acquisition of temporality. In C. Perdue (Ed.), *Adult Second Language Acquisition: Cross-linguistic perspectives,* Vol. 2: *The results* (pp. 73–118). Cambridge: CUP.

Klein, W. (1995). The acquisition of English. In R. Dietrich, W. Klein, & C. Noyau (Eds.), *The Acquisition of Temporality in a Second Language* [Studies in Biliingualism 7] (pp. 31–70). Amsterdam: John Benjamins.

Klein, W., Dietrich, R., & Noyau, C. (1995). Conclusions. In R. Dietrich, W. Klein, & C. Noyau (Eds.), *The Acquisition of Temporality in a Second Language* [Studies in Bilingualism 7] (pp. 261–280). Amsterdam: John Benjamins.

Klein, W. & Clive, P. (1997). The basic variety (or: Couldn't natural languages be much simpler?). *Second Language Research, 13*, 301–347.

Lalleman, J. (1996). The state of the art in second language acquisition research. In P. Jordens & J. Lalleman (Eds.), *Investigating Second Language Acquisition* (pp. 3–69). Berlin: Mouton de Gruyter.

Lefebvre, C. (1996). The tense, mood, and aspect system of Haitian Creole and the problem of transmission of grammar in creole genesis. *Journal of Pidgin and Creole Languages, 11*, 231–311.

Lefebvre, C. (1998). *Creole Genesis and the Acquisition of Grammar: The case of Haitian Creole*. Cambridge: CUP.

Lefebvre, C. & Lumsden, J. S. (1994). Relexification in creole genesis. Paper read at the MIT Symposium on the role of relexification in creole genesis: The case of Haitian Creole.

Lumsden, J. S. (1999). Language acquisition and creolization. In M. DeGraff (Ed.), *Language Creation and Language Change: Creolization, diachrony and development* (pp. 129–157). Cambridge MA: The MIT Press.

Mather, P.-A. (2000). Creole genesis: Evidence from West African L2 French. In D. G. Gilbers, J. Nerbonne, & J. Schaeken (Eds.), *Languages in Contact* (pp. 247–261). Amsterdam: Rodopi.

McWhorter, J. H. (1999). A Creole by any other name: Streamlining the terminology. In M. Huber & M. Parkvall (Eds.), *Spreading the word: The issue of diffusion among the Atlantic creoles* (pp. 5–28). London: University of Westminster Press.

Migge, B. (1998). Substrate influence in creole formation: The origin of *give*-type serial verb constructions in the Surinamese Plantation Creole. *Journal of Pidgin and Creole Languages, 13*, 215–265.

Migge, B. (2004). Tracing the origin of modality in the creoles of Suriname. Ms, University College, Ireland.

Migge, B. & Winford, D. (2003). Gbe substrate influence on the TMA systems of the Surinamese Creoles. Paper presented at the workshop on Transatlantic Sprachbund, Amsterdam, April 2003. [To appear in *Journal of Pidgin and Creole Languages*.]

Mufwene, S. (1990). Transfer and the substrate hypothesis in creolistics. *Studies in Second Language Acquisition, 12*, 1–23.

Mufwene, S. (1996a). The development of American Englishes: Some questions from a creole genesis perspective. In E. W. Schneider (Ed.), *Focus on the* USA [Varieties of English Around the World. G16] (pp. 231–264). Amsterdam: John Benjamins.

Mufwene, S. (1996b). The Founder Principle in creole genesis. *Diachronica, 13*, 83–134.

Mufwene, S. (2001). *The ecology of language evolution*. Cambridge: CUP.

Neumann-Holzschuh, I. & Schneider, E. W. (Eds.) (2000). *Degrees of Restructuring in Creole Languages*. Amsterdam: John Benjamins.

Patterson, O. (1982). *Slavery and Social Death*. Cambridge MA: Harvard University Press.

Postma, J. M. (1990). *The Dutch in the Atlantic Slave Trade: 1600–1815*. Cambridge: CUP.

Roberts, S. J. (1998). The role of diffusion in the genesis of Hawaiian Creole. *Language, 74*(1), 1–39.

Roberts, S. J. (2000). Nativization and the genesis of Hawaiian Creole. In J. McWhorter (Ed.), *Language Change and Language Contact in Pidgins and Creoles* (pp. 257–300). Amsterdam: John Benjamins.

Schlyter, S. (1990). The acquisition of French temporal morphemes in adults and in bilingual children. In G. Bernini & Giacalone Ramat (Eds.), *La temporalità nell'acquisione di lingue seconde* (pp. 293–308). Milan: Franco Angeli.

Schumann, C. L. (1783). Neger-Englishes Wörterbuch. MS. In A. Kramp 1983, Early creole lexicography. A study of C. L. Schumann's manuscript dictionary of Sranan (pp. 44–305). Ph.D dissertation, University of Leiden.

Seuren, P. (1981). Tense and aspect in Sranan. *Linguistics, 19*, 1043–1076.

Siegel, J. (1999). Transfer constraints and substrate influence in Melanesian Pidgin. *Journal of Pidgin and Creole Languages, 14*, 1–44.

Siegel, J. (2000). Substrate influence in Hawai'i Creole English. *Language in Society, 29*, 197–236.

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*, 319–333.

Singler, J. V. (1993). African influence upon Afro-American language varieties: A consideration of sociohistorical factors. In S. Mufwene (Ed.), *Africanisms in Afro-American language varieties* (pp. 235–253). Athens GA: University of Georgia Press.

Singler, J. (1995). The demographics of creole genesis in the Caribbean: A comparison of Martinique and Haiti. In J. Arends (Ed.), *The Early stages of Creolization* [Creole Language Library 13] (pp. 203–232). Amsterdam: John Benjamins.

Singler, J. (1996). 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.

Spears, A. (1990). Tense, mood and aspect in the Haitian Creole preverbal marker system. In J. Singler (Ed.), *Pidgin and Creole Tense-Mood-Aspect Systems* [Creole Language Library 2] (pp. 119–142). Amsterdam: John Benjamins.

van Buren, P. (1996). Are there principles of universal grammar that do not apply to second language acquisition? In P. Jordens & J. Lalleman (Eds.), *Investigating Second Language Acquisition* (pp. 187–207). Berlin: Mouton de Gruyter.

van den Berg, M. (2000). *Mi no sal tron tongo Early Sranan in court records 1667–1767*. MA Thesis, Katholieke Universiteit Nijmegen.

van den Berg, M. (2004). Tense, mood and aspect in early Sranan. Ms.

van Dyk. n.d. [c1765] *Nieuwe en Nooit Bevoorens Geziene Onderwyzinge in het Bastert Engels, of Neeger Engels, zoo als het Zelve in de Hollandsze Colonien Gebruikt Word*. Amsterdam: Jacobus van Egmont.

Voorhoeve, J. (1957). The verbal system of Sranan. *Lingua, 6*, 374–396.

Voorhoeve, J. & Lichtveld, U. M. (1975). *Creole Drum: An anthology of creole literature in Suriname*. New Haven CT: Yale University Press.

Winford, D. (2000a). Tense and aspect in Sranan and the creole prototype. In J. McWhorther (Ed.), *Language Change and Language Contact in Pidgins and Creoles* [Creole Language Library 21] (pp. 383–442). Amsterdam: John Benjamins.

Winford, D. (2000b). Intermediate creoles and degrees of change in creole formation: The case of Bajan. In I. Neumann-Holzschuh & E. W. Schneider (Eds.), *Degrees of Restructuring in Creole Languages* [Creole Language Library 22] (pp. 215–246). Amsterdam: John Benjamins.

Winford, D. (2002). Creoles in the context of Contact Linguistics. In G. Gilbert (Ed.), *Pidgin and Creole Linguistics in the twenty-first century* (pp. 287–354). New York: Peter Lang.

Syntactic properties of negation in Chinook Jargon, with a comparison to two source languages

Zvjezdana Vrzić
New York University, USA

This paper analyzes the syntactic features of sentential negation in Chinook Jargon (CJ) and compares them with the features of Lower Chinook and Upper Chehalis. The analyses is carried out within the framework of Principles and Parameters syntax and looks at the positioning and the categorical status of the CJ negative marker, followed by the analysis of the relationship of the negative marker with negative NPs within the clause. The comparison of sentential negation in CJ and its source languages establishes that important similarities exist between CJ and its source languages, suggesting that these properties in CJ are due to contact influence. The paper ends with a brief discussion of how these findings reflect on the issue of linguistic constraints on contact influence.

0. Introduction[1]

Chinook Jargon (CJ) is a pidgin that arose through the contact among several American Indian languages, French, and English in Oregon in the United States around the beginning of the 19th century (Thomason 1983; Hajda, Zenk, & Boyd 1988).[2] It was used widely as a *lingua franca* in the course of the 19th century and at the beginning

1. I would like to thank two anonymous reviewers for their valuable comments and corrections to this paper. I have also benefited from the feedback provided to me by the audience at the 34th International Conference on Salish and Neighboring Languages in Kamloops, British Columbia, where an earlier version of this paper was presented. All mistakes and imperfections are my own.

2. The exact timing of the formation of CJ is an issue that has not been settled yet in the literature. The debate revolves around the question of whether CJ was formed before or after the contact of the indigenous population of Oregon with Europeans and Euro-Americans (see Thomason 1983). See references in the text and Vrzić (1999) for a discussion of the various 'scenarios' of CJ origin.

of the 20th century in Oregon, Washington, British Columbia and southern Alaska in Indian-Indian and white-Indian communication.

The purpose of this paper is to analyze the syntactic properties of sentential negation in CJ.[3] First, I address the issue of the positioning of the negative marker, its categorial status and the relationship of the negative marker with negative NPs within the clause. Then, these features of sentential negation in CJ are compared to the syntactic characteristics of sentential negation in Lower Chinook (Chinookan) and Upper Chehalis (Salishan), two languages considered to be the most relevant CJ source languages. I conclude that significant similarities exist between the properties of the negative construction in CJ and the two source languages, suggesting that the properties of negation in CJ are due to transfer and contact influence. The paper ends with a short discussion of how these findings about the syntax of negation in CJ reflect on the process of pidgin/creole genesis and the issue of linguistic constraints on contact influence.

1. Syntactic properties of CJ sentential negation

1.1 Positioning of the negative marker

The lexical form of the CJ negative marker is *wek*.[4] The examples in (1) below show that the negative marker *wek*, underlined in examples (1a, b), precedes the pronominal subject, which is then followed by the verb and the object in (1a) and the non-verbal predicate in (1b). This word order is without exception.[5]

(1) a. *Pi* *wek* *nsaika* *nanich* *yaka.*[6]
and NEG 1PL see 3SG
'And we didn't see him'

3. The syntactic analysis of CJ negation is based on a portion of CJ texts that I have compiled, transcribed from Duployan shorthand, and translated (Vrzić 1999; Vrzić forthcoming). These texts were originally published in the late 19th and early 20th centuries in a publication *Kamloops Wawa* in British Columbia.

4. CJ negative marker *wek* was borrowed from Nootka, one of the Native American languages of the Pacific Northwest that played a role in its formation.

5. The corpus used in this paper consists of 109 negative sentences with *wek*. The total number of negative sentences found in the corpus used in Vrzić (1999) is 181: 13 sentences use a negative adverb *wek kansih* (*weht*) 'never (again)', 46 use a negative modal expression *wek kata* (*weht*), and 13 use a negative marker *elo* 'not', which are not discussed in this paper.

6. The following abbreviations will be used in glossing the examples and in the text:

1SG, etc. 'first person singular personal pronoun'
1PL, etc. 'first person plural personal pronoun'

b. *Wek yaka kwash.*
NEG 3SG afraid
'He was not afraid'

The examples in (2) below show that the same word order appears in embedded sentences. In (2a) *wek* is followed by a pronominal subject, which is in turn followed by a verbal predicate. In (2b) *wek* is followed by a pronominal subject and a non-verbal predicate.

(2) a. *Ikta alta yaka mamuk kopa msaika pus wek msaika kopet mamuk masachi.*
what now 3SG make PREP 2PL C NEG 2PL finish make sin
'What did he do to you so that you do not stop doing bad things?'

b. *Tlus msaika kwanesem tlus nanich pus wek msaika tsepe.*
good 2PL always good see C NEG 2PL mistaken
'You should always watch so that you are not mistaken'

The fact that pronominal subjects are placed outside the VP is confirmed in (3) below where the pronominal subject is separated from the verb by a VP-adverb (in bold).

(3) *Wek kata nsaika* ***ayak*** *kuli.*
NEG how 1PL fast go
'We couldn't go fast'

Based on the data presented in (1), (2) and (3), it can be concluded that the word order pattern for matrix and embedded negative clauses containing pronominal subjects in CJ is as in (4) below.

(4) (C) NEG S_{pro} (Adverb) V (O)
AP
NP

Examples including nominal subjects are discussed next. Note that a nominal subject (in brackets) can either precede or follow the negative marker *wek,* as shown in (5a) and (5b) respectively.[7]

NEG 'negation'
C 'complementizer'
PREP 'preposition'
DEM 'demonstrative'
S_{nom} 'nominal subject'
S_{pron} 'pronominal subject'

7. The number of sentences with nominal subjects not followed by pleonastic pronouns was much smaller than those with pronominal subjects and those with both nominal and pleonastic

(5) a. [*Kopet iht Noe kanamokst yaka tanas*] *wek memlus.*
only one Noah with 3SG child NEG die
'Only Noah and his children didn't die'

b. *Pi wek* [*S.T.*] *mash komtaks Noe.*
but NEG God leave know Noah
'But God didn't forget Noah'

The same facts are observed in embedded sentences: In (6a) the nominal subject follows *wek* and in (6b) the nominal subject precedes it.

(6) a. *Yaka mamuk mitlait iht lesash kopa ukuk tlus elehe yaka port*
he make stay one angel PREP DEM good land 3SG door
pus wek [*Adam pi Ev*] ***weht*** *kilapai kopa yaka.*
C NEG Adam and Eve again return PREP 3SG
'He placed an angel on heaven's door so that Adam and Eve don't return to it again'

b. *Ayu naika mamuk pus* [*masachi*] *wek tolo naika.*
much 1SG make C sin NEG win 1SG
'I try hard so that sin does not win over me'

As in the case of a pronominal subject (see (3) above), a nominal subject following *wek* can be separated from the verb by an adverb. In (6a) above, the adverb *weht* 'again' (in bold) precedes the verb. In (7) below the subject NP is separated from the verb by the adverb *dlet* (in bold).

(7) *Yaka ukuk pus wek S.T.* ***dlet*** *mas nsaika.*
3SG DEM C NEG God truly leave 1PL
'This is so because God didn't truly abandon us'

If the nominal subject is accompanied by a pleonastic pronominal subject, the nominal subject precedes *wek* as shown in (8a), which has a verbal predicate, and in (8b), which has a non-verbal predicate.[8]

(8) a. *Pi* [*Maik*] *wek yaka lolo yaka kiutan klaska lains kopa yaka*
but Mike NEG 3SG carry 3SG horse 3PL reins PREP 3SG
lema.
hands
'However, Mike didn't hold his horse's reins in his hands'

---

pronominal subjects. Claims made in this paper regarding these structures would need to be reconfirmed on a larger corpus.

8. Pleonastic subjects have been claimed to be quite common in the Oregon variety of CJ (see Jacobs 1936; Thomason 1983; Zenk 1984). This feature is less dramatically present in the British Columbia variety of CJ under consideration here (see Vrzić 1999).

b. [*Ukuk iht person*] *wek yaka ukuk weht iht.*
DEM one person NEG 3SG DEM again one
'The first person is not the other one'

In summary, the following word order patterns, laid out in (9), are possible in CJ based on the data presented in (1) through (8). Note again that a pronominal subject never precedes the negative marker.

(9)

	NEG	S_{pro}	Predicate
	NEG	S_{nom}	Predicate
S_{nom}	NEG	S_{pro}	Predicate
S_{nom}	NEG		Predicate
$*S_{pro}$	NEG		Predicate

The positioning of adverbs with respect to pronominal and nominal subjects shows that overt subjects in CJ are outside the VP (see examples (6a) and (7) above), presumably, in [Spec, AgrSP], the canonical surface position of derived subject (see Chomsky 1993). It follows then that the negative marker *wek* is in a pre-IP position or in a position between the left edge of the AgrSP and the complementizer *pus* in CP:

(10) $[_{CP}$ *pus* [*wek* $[_{AGRSP}$ $S_{pro/nom}$ [Adv $[_{VP}$ …]]

When nominal subjects precedes *wek,* as in (5a), (6b), and (8a, b) above, they must be in the Spec position of one of the functional projections within the CP. Assuming the expanded, articulated structure of the CP proposed by Rizzi (1997: 297), presented in (11) below, the nominal subject in CJ could be located in either the Top(ic)P(hrase) or in the Foc(us)P(hrase).

(11) [ForceP TopP* FocP TopP* FinP $[_{IP}$ …]]
└──────────────────────┘
Expanded CP

Keeping the usual syntactic facts about topic and focus NPs in mind (see Rizzi (1997) and references within), the topic analysis is plausible for examples in which a nominal subject is combined with a pleonastic pronominal subject (as in (8a, b)), and the focus analysis may be applicable to a preposed nominal subject not followed by a pleonastic pronominal (as in (5a)).[9]

9. While this paper does not discuss such structures specifically, an analysis assuming that NPs preceding *wek* are either Topic NPs or Focus NPs seems plausible. The placement of nominal subjects followed by pleonastic pronominal subjects such those in (8a, b) is likely to be related to a change in discourse topic. On the other hand, preposed nominal subjects not followed by pleonastic pronominal subjects such as those in (5a, 6b) are likely to be related to sentential focus. While the exact function of these NPs could be confirmed only through the analysis of their use in connected speech or text, as suggested by an anonymous reviewer of this paper, the assumptions made about their syntactic function, based on the word order analysis presented here, are supported by the fact that examples also exist where nominal subjects follow *wek* and

1.2 The categorial status of the CJ negative marker 'wek'

The discussion in 2.1 did not touch on the issue of the categorial status of *wek*. The issue is whether *wek* is an adverb or adverbial particle in an adjoined position or a functional element with its own projection, NegP. In the latter case, the question is also whether *wek* is a head or a specifier of this projection. Both of these questions are addressed in this section.

Since *wek* in CJ has a fixed position, as shown in 2.1, it seems justified to consider *wek* a particle, i.e. a free functional element with its own projection, rather than an adverb adjoining to other phrases. While the position of most adverbs in CJ is generally fixed, variation is possible. For instance, temporal adverbs (e.g., *alta* 'now', *alke* 'later, after', *ankante* 'earlier, before') usually precede the IP, but these adverbs can also occur in sentence-final or sentence-medial positions (i.e., between the subject and the verb). In contrast, no variation in the positioning of the negative marker *wek* is found, supporting the claim that *wek* is within its own projection, NegP. Such representation of negative particles is unsurprising and has become very common for many non-pidgin/creole languages since Pollock (1989).

Regarding the position of the NegP in CJ, the facts discussed in 2.1 suggest that it is a functional projection immediately dominating the IP, within the expanded CP (see (11) above).[10] Hence, *wek* is an element that has a complementizer-like status in CJ. This property of CJ negation is not unusual. There is ample evidence for the presence of negative constituents in this position in the clausal structure from the analysis of non-pidgin/creole languages such as Latin, Celtic, and Basque.[11] Rizzi (1997) does

---

are presumably in AgrSP position, such as in examples (5b, 6a, 7). This analysis of the patterns in (9) is also desirable because it promises to explain an alternation that would otherwise remain an unexplained variation.
The plausibility of this analysis is illustrated with the following example from the Bible History (KW, III:9:154). The relevant sentence, involving an NP preposed for focus, is in italics.

> Kakwa kanawe telikom memlus kanamokst kanawe hloima mawich: ber, [...], kanawe kalakala pi kanawe ikta kuli kopa ukuk elehe. [*Kopet iht Noe kanamokst yaka tanas*] <u>*wek*</u> *memlus pi kanawe ikta mitlait kopa yaka ayas knim wek memlus.*
> 'Hence, all the people died together with all different animals: bear, [...], all birds and everything that runs on the ground. *It was Noah alone with his children who didn't die, and it was everything that stayed on his big boat that didn't die.*'

10. I will not distinguish between FinP and IP since the issue of the expression of finiteness in CJ is not discussed in this paper.

11. For example, Laka (1990) suggests that Basque has a ΣP (Speech Act Phrase) containing features [± negative]. The [+ negative] feature is instantiated by a negative complementizer. Culicover (1991, 1993) has posited the existence of PolP (Polarity Phrase) in English negative adverb and PP fronting constructions. See further references given by Rizzi (1997) and Haegeman (1995) for a discussions of cross-linguistic issues in the syntax of negation. Evidence for CP-iteration or Double-CP constructions (not related to negation) was first given by Platzak

not specifically address the issue of negation within the CP system, perhaps because Italian, the language he discusses, has IP-internal negation. However, he does mention that negation, like mood, agreement, or tense, is one of the syntactic features, normally associated with and expressed within the IP-system, that can be 'replicated' (i.e., also expressed, often redundantly) in the complementizer system, usually using free morphemes rather than affixation.

In summary, based on the discussion above, CJ clausal structure posited in (10) can be further articulated as shown in (12) below. The schema in (12) shows that the expanded CP in CJ consists of at least three functional projections: the CP (in the narrow sense), the optionally projecting TopP/FocP,[12] and the NegP.

(12) *CJ clause structure with the expanded CP*
$[_{\text{CP}}$ *pus* $[_{\text{(TopP/FocP)}}$ S_{nom} $[_{\text{NegP}}$ *wek* $[_{\text{AgrSP}}$ $\text{S}_{\text{pro/nom}}$ [Adv $[_{\text{VP}}$ …]]

A further question, discussed below, is whether *wek* is a head or a specifier of NegP. This question is related to the issue of whether CJ has negative concord, which is addressed first.

Negative concord is a phenomenon which requires that, when a sentential negative marker is present in a sentence, all indefinite elements in the sentence be also marked as negative. Hence, negative concord is a kind of an 'agreement' process (see Zanuttini 1991): The meaning of the sentence does not reflect the occurrence of multiple negative items, since the sentence is interpreted as a simple negative statement. This can be illustrated with an example from Croatian/Serbian, which, like many other languages, has negative concord. While multiple elements in the sentence in (13) are negative, the meaning of the sentence is simple negation, as shown by the English translation.

(13) *Nitko* *nikada* *nije* *ništa* *nikome* *rekao.*
Nobody-NOM never hasn't nothing nobody-DAT said
'Nobody has ever said anything to anybody'

Standard English, on the other hand, does not have negative concord and only one negative item per sentence is allowed if the interpretation of the sentence is to be negative, see (14a) below. When two negative elements are present in a Standard English sentence, they 'cancel' each other out and the interpretation of the sentence is positive, see (14b) below.[13]

(14) a. I didn't see anything = I saw nothing
b. I didn't see nobody = I saw somebody

(1986) for Swedish double complementizer constructions. Double-CP constructions also exist in Croatian/Serbian, see Vrzić 1996.

12. I did not find CJ examples containing more than one TopP or both a TopP and a FocP.

13. These facts are analyzed in great detail in the rich literature on the syntax of negation, see for example, Haegeman (1995) and Zanuttini (1997), as well as references given therein.

Like English, CJ does not seem to have negative concord. The presence of the sentential negative marker *wek* is in complementary distribution with negative indefinites (i.e., negative quantifiers like *elo ikta* 'nothing', *elo klaksta* 'nobody', etc.), as shown in examples in (15) in contrast to those in (16) below.[14]

(15) a. *Elo klaksta mamuk kopa ukuk senmokst son.*
NEG anybody work on DEM seven day
'Nobody works on the seventh day'

b. *Elo ikta masachi naika mamuk.*
NEG anything bad 1SG make
'I did nothing bad'

c. *Elo ikta yaka eskom.*
NEG anything 3SG take
'He took nothing'

In (15a, b, c) the negative marker *wek* is absent, but the presence of negative quantifiers ensures that the sentences is interpreted as a negative statement. In the examples in (16), the negative marker *wek* is used, and the indefinite words (in bold) show up in their positive form, acting as negative polarity items, similar to *anybody*, *anything*, etc. in English.[15]

(16) a. *Pi wek klaska wawa* ***ikta*** *kopa yaka.*
and NEG 3PL say anything PREP 3SG
'And I didn't tell him anything'

b. *Wek kata nsaika mamuk* ***ikta.***[16]
NEG anyhow 1PL do anything
'We couldn't do anything'

---

14. Several negative words use a different constituent negative marker – *elo*, e.g., *elo ikta, elo klaksta, elo ayu*. Used alone *elo* also means 'none, nothing', as in the expression *chako elo* 'become nothing; disappear'. In a few instances, *elo* also seems to be used to mark sentential negation:

(i) *Kakwa elo yaka mas pepa kopa maika.*
SO NEG 3SG send paper PREP 2SG
'So, he didn't send a paper to you'

Examples of negative sentences involving *elo* represent a small percentage of the total sample as noted earlier (see Note 5). Most of these are occurrences of a negative indefinite like 'nobody' and 'nothing.' Note that *elo* functioning as a sentential negative marker is not known in Oregon CJ.

15. CJ negative polarity items, such as *klaksta* 'anybody', *ikta* 'anything', *kah* 'anywhere', etc., are also used as interrogatives (e.g., *klaksta?* 'who') and indefinites (e.g., *klaksta* 'somebody').

16. The use of the marker *wek kata* instead of *wek* alone in this and other examples will be discussed below.

c. *Wek* *kata* *weht* *yaka* *stop* ***kah.***
NEG anyhow again 3SG stop anywhere
'Again, he couldn't stop anywhere'

d. *Wek* *kata* *naika* *saliks* *kopa* ***klaksta.***
NEG anyhow 1PL angry PREP 3PL
'I couldn't be angry with anybody'

Now that it is established that CJ does not allow negative concord, the issue of whether *wek* is a head or a specifier of the NegP projection can be addressed. It has been proposed (Rizzi 1996 [1991]; Zanuttini 1991; Haegeman & Zanuttini 1991, among others) that the behavior of negative elements can be likened to the behavior of WH-words, described by the WH-Criterion. Namely, in one interpretation (proposed by Rizzi 1996 [1991]), a WH-phrase needs to enter into a Spec-Head relationship with a head carrying a [+wh] feature, and each WH-head must be in a Spec-Head relation with a WH-phrase. This condition on the licensing of WH-phrases, which parallels the licensing requirements of referential NPs moving for agreement, was named the WH-Criterion (expressed here informally, see Rizzi 1996 [1991] and Haegeman & Zanuttini 1991 for details). This criterion was meant to explain the occurrence of WH-movement and auxiliary raising/do-support in English WH-questions. It is taken to be satisfied at the level of Logical Form (LF) in the languages where WH-phrases occur in base-generated positions.

Haegeman and Zanuttini (1991), among others, follow the same logic in their analysis of negative words in languages which allow negative concord, such as West Flemish, in the form of the Neg-Criterion. In West Flemish, when a negative word, such as *niemand* 'nobody', precedes the negative marker *nie*, a sentence is interpreted as a simple negative statement and stands as an example of negative concord in West Flemish, see (17a) below. In contrast, as (17b) shows, when *niemand* follows *nie*, the sentence is interpreted as involving double negation, just like its English translation.

(17) a. … *da* *Valere* *niemand* *nie* *kent.*
that Valere nobody not knows
'… that Valere does not know anybody'

b. … *da* *Valere* *nie* *niemand* *kent.*
that Valere not nobody knows
'… that Valere doesn't know nobody'

Haegeman and Zanuttini (1991) propose that in (17a) the negative word *niemand* moves to an adjoined position in the clause, higher than the VP, in order to check its features, thus enabling the negative concord reading of the sentence (see Haegeman & Zanuttini 1991; Haegeman 1995; Haegeman & Zanuttini 1996 for further details).

DeGraff (1993: 67), following Zanuttini (1991), applies a slightly different but related analysis to Haitian, where negative sentences with two negative elements have a negative concord reading:[17]

(18) a. *Pèsonn* *pa* *vini.*
nobody *pa* come
'Nobody has come.'
b. *Mwen* *pa* *wè* *pèsonn.*
1SG *pa* see nobody
I haven't seen anybody.

He claims that the negative XP in (18a) is in a [Spec, NegP] position. If negative concord is a type of agreement process, as proposed by Zanuttini (1991), then the negative indefinite needs to raise to the [Spec, NegP] either at S-structure or LF to check its features with the head of NegP.[18] If the [Spec, NegP] were filled with the negative marker, the raising of the negative word would be blocked and negative concord in a language like Haitian would be impossible. Therefore, the negative marker *pa* in Haitian must be a head of the NegP phrase, rather than an XP.

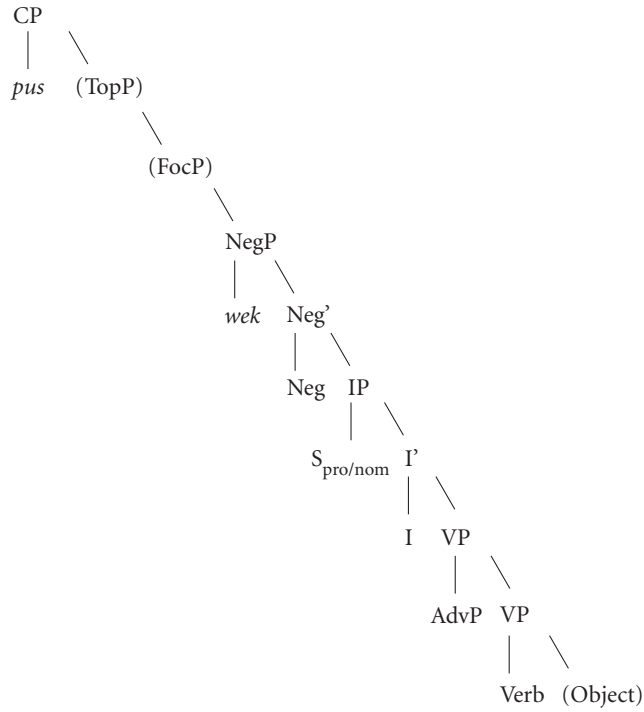
Consequently, if the same logic is applied to CJ, the fact that it lacks negative concord, illustrated in examples (15) and (16) above, suggests that the negative marker *wek* may be located in [Spec, NegP], and hence, that it is a phrasal constituent.[19] The structure of the CJ clause given in (12) above can now be further refined as follows (irrelevant structure is omitted):

17. Note that the negative properties of *pèsonn* are confirmed by the fact that it means 'nobody' in isolation.

18. Judging from the word order in (18b), the raising of the negative quantifier is covert, that is, it happens at LF in Haitian, unlike in West Flemish, see (17).

19. As suggested by an anonymous reviewer this evidence is plausible, but not conclusive: The correlation between the presence of negative concord and the head status of the negative marker works for some languages, like Haitian, but not for others, like West Flemish and Bavarian, where negative markers are phrasal elements. However, the facts presented in Section 2.4 provide some further support for the phrasal status of *wek*: Negative object NPs must obligatorily precede the subject and the verb in CJ. If they raise overtly to [Spec, NegP], as seems plausible, their complementary distribution with *wek* can be explained that by the fact that both negative NPs and *wek* are phrasal elements 'competing' for the same position in NegP.

(19)



1.3 'Complex markers' of sentential negation in CJ

CJ has a number of negative expressions that consist of the negative marker *wek* and another lexical element, such as *wek kansih* (*weht*), *wek kata* (*weht*), *wek kata pus*, *wek tlus pus*, *wek saya*, *wek lili*, etc. The issue of whether these negative constituents represent local or sentential negation will be addressed in this section. The following examples illustrate each of these negative expressions (underlined).

(20) a. *Wek kansih naika mamuk taye ukuk kluchmin.*
NEG how-much 1SG make chief DEM woman
'I will never bow to this woman'

b. *Wek kansih weht naika kilapai kopa masachi.*
NEG how-much again 1SG return PREP sin
'Never again will I return to sin'

c. *Pi wek kata nsaika mamuk ikta.*
and NEG how 1PL do what
'And we couldn't do anything'

d. *Pi wek kata weht yaka stop kah.*
and NEG how again 3SG stop where
'And he couldn't stop anywhere anymore'

e. *K'el yaka latet Ken, wek kata pus yaka kopet saliks.*
hard 3SG head Cain NEG how C 3SG stop angry
'Cain was stubborn; he couldn't stop being angry'

f. *Wek kata pus ukuk pichon tlap dlai elehe, kakwa ayak yaka kilapai.*
NEG how C DEM pigeon get dry land so quickly 3SG return
'The pigeon couldn't reach the dry land, so it quickly returned'

g. *Wek tlus pus nsaika mamuk memlus nsaika ow.*
NEG good C 1PL make die 1PL brother
'We shouldn't kill our brother/s'
(Lit. 'It is not good for us to kill our brother/s')

h. *Wek saya pulakli, kakwa nsaika mash ukuk ayas stiwil hows.*
NEG far night so 1PL leave DEM big prayer house
'It was close to the night, so we left the church'

i. *Wek lele S.T. mamuk paya yaka kanamokst telikom.*
NEG long.time God make fire 3SG with people
'Soon God will burn it together with the people'

Judging from the interpretation of the sentences (20a) through (20f), the complex negative expressions under consideration affect the polarity of the whole sentence and therefore, function as negative operators, taking a whole sentence within their scope. In contrast, in examples (20h) and (20i), the negation is strictly local and affects only the meaning of the lexical item it modifies. The effect of negative operators (often adverbial elements such as English 'never') has been attested in various languages (see Culicover 1991, 1993; Haegeman 1995, and references therein).

Unlike preposed negative operators in English, CJ negative operators do not trigger auxiliary inversion simply because CJ does not seem to have verb movement of any kind (see Vrzić 1997, 1998, 1999). It is not possible to establish whether negative operators in matrix clauses block the extraction of WH-phrases from embedded clauses as would be expected. No long extraction of WH-phrases is attested in CJ with or without negation. However, clear evidence does exist that (20a–f) indeed involve negative operators. Namely, examples (20c) and (20d) contain items *ikta* 'what' and *kah* 'where' whose interpretation as negative polarity items ('anything' and 'anywhere' respectively) is licensed by the existence of negative operators.

With respect to the positioning of negative operators, it seems reasonable to assume that they are base-generated in the same position as the negative marker *wek*, i.e. in [Spec, NegP] (see (19) above). Unlike English 'never,' for instance, CJ *wek kansih* (*weht*) is always found before the subject and the VP, just like the negative marker *wek*.[20]

20. With regard to *wek kata* (*weht*), as its English translation suggests (see 20c, d), it is a kind of negative modal particle. It would be tempting to say that *wek kata* (*weht*), like other neg-

1.4 Positioning of negative indefinites in CJ

Zanuttini (1994, see also 1997) proposes that in different languages negative markers can be base-generated in different positions in the clause, and that languages can have several functional categories related to negation. One of the positions, the highest in the structure of a clause and obligatorily dominating the T(ense)P, is labeled PolP in Zanuttini (1994) or NegP-1 in Zanuttini (1997). NegP-1 is the position where the negative marker moves, either overtly or covertly depending on the strength of its features, so that scopal relations can be interpreted. Zanuttini (1997: 11) further proposes a typology of negative marking: "NegP-1 has 'strong' features in the languages that express sentential negation by means of a pre-verbal negative marker which by itself can negate the clause, and 'weak' features in the languages that express sentential negation by means of a negative marker of another kind".

CJ, which has a single, pre-verbal negative marker, is expected to have a NegP (equaling Zanuttini's NegP-1) with strong features, according to Zanuttini's typology. Strong features of NegP need to be 'checked off' through overt syntactic movement. This checking requirement seems to be trivially satisfied in CJ in negative sentences containing negative marker *wek* (or others mentioned in 1.3 above), because, as argued above, *wek* is base-generated in NegP. However, when negative indefinites are present in a CJ clause, because CJ does not have negative concord and negative indefinites function as negative quantifiers, i.e. operators, the existence of the sentential negative marker *wek* is precluded. The expectation that NegP should have strong features in CJ, to be checked overtly by an appropriate element, is confirmed by the positioning of negative quantifiers, in particular, when they are syntactic objects. Objects normally follow verbs in CJ. However, as shown in (21), the object negative quantifier is preposed, presumably raised overtly to [Spec, NegP] in satisfaction of the checking requirements of a negative head with strong features.[21]

(21) a. *Elo* *ikta* *masachi* *naika* *mamuk* [...]
NEG what bad-thing 1PL do
'I did nothing wrong...'

ative operators, is located in a functional category which includes both negative and modal features, i.e., a PolP. Ample evidence exists for such a projection, with either negative or modal properties, across languages. In addition to references mentioned earlier for negation, see also Dobrovie-Sorin (1994), Rudin (1985a), (1985b), Rivero (1994), among others. Alternatively, as an anonymous reviewer suggested, *wek kata* (and CJ negative operators) could be located in the area of the clausal structure where certain modal projections are present as well. Deciding on this issue would require looking into the properties of *wek kata* and other modal expressions in CJ more closely, which is beyond the scope of this paper.

21. As suggested by an anonymous reviewer, this data could be taken to support claims made by Kayne (1998), specifically, a claim that negative quantifiers must raise overtly to NegP.

b. *Elo* *ikta* *yaka* *eskom.*
NEG what 3SG take
'He took nothing'

Compare the examples in (21) with the closely parallel example given in (22) below, where the negative marker *wek* is present and *ikta* 'anything' (in bold), the positive indefinite NP object (i.e. the negative polarity item) remains in its original position.

(22) *Pi* *wek* *klaska* *wawa* ***ikta*** *kopa* *yaka.*
and NEG 3PL say what PREP 3SG
'And they didn't tell him anything'

In summary, a negative sentence with an indefinite NP object can be expressed in CJ either by a) preposing a negative quantifier *elo ikta* (see (21)), or b) by using a negative polarity item *ikta* in a sentence introduced by the negative marker *wek* (see (22)).[22]

1.5 Summary of the section

Several syntactic properties of CJ negation have been established in this section: a) the negative marker *wek* is located in a pre-IP position; b) the negative marker *wek* is a Specifier of the NegP, the phrase it is generated in; c) CJ doesn't allow negative concord; d) CJ has other negative operators, e.g., *wek kansih* (*weht*) and *wek kata* (*weht*), which are generated in the same position as *wek;* e) Negative quantifiers raise overtly in CJ to check off the strong features of NegP. When a negative quantifier is an object NP, this results in the non-canonical, OSV word order.

2. Negation in CJ source languages

In this section, the basic properties of sentential negation in two important source languages, Lower Chinook and Upper Chehalis, will be discussed. The positioning of

---

22. There seems to be another, much less common, way of expressing a similar meaning, see (i) below.

(i) *Yaka* *elo* *tomtom* *ikta.*
3SG NEG think what
'He didn't worry about anything'.

Interestingly, neither of the following ways of expressing the similar meaning is attested in the texts inspected:

(ii) Subject Verb *elo ikta* (the negative quantifier doesn't move)
(iii) *elo* Subject Verb *ikta* (*elo* is in a pre-IP position)

As noted earlier, *elo* is quite rare as a marker of negation, and I set aside its description in this paper.

the negative marker in these languages can be made out with considerable certainty. With regard to other properties, since no detailed analyses or data are available, I can only suggest what these might likely be, based on my own, cursory look at the available sources. It should be noted that, in addition to the two languages discussed here, Thomason (1983:855) gives examples from various other American Indian languages of the Pacific Northwest. She shows that all have 'sentence-initial negatives' which can be either particles, as in Lower Chinook, or intransitive verbs, as in Nootka.

2.1 Lower Chinook

The Lower Chinook negation marker was *nikšt* (*nēkct* in Boas' (1910/11) spelling), a free morpheme. Boas classifies this negative marker as an adverb and a particle, i.e., a non-inflected lexical form, see examples (23a, b) below (p. 668).[23] It is clear from the examples, that the negative marker (underlined) precedes the verb (see (23a)) or the attribute complement (see (23b)).[24]

(23) a. [...] *nēkct L!pax aqLāx.*[25]
not well someone.makes.him
'He is not made well'

b. [...] *a'lta nēkct qa'nsix t!ayā' aqLāx.*
now not (any)how well someone.makes.him
'He cannot be made well at all'

Boas does not discuss further the properties of negation in Lower Chinook, including the structure and function of negative words. However, the example (23b) contains the negative operator *nēkct qa'nsix*, which consists of the negative marker *nēkct* and the question word *qa'nsix,* quite similar to CJ. In his discussion of Wishram (an Upper Chinook dialect), Dyk (1933) is equally brief on negative particles. He mentions two of them – *k'aya* 'no' and *naqi* 'not', and gives the following example, where the negative marker (underlined) also precedes the verb.

(24) *K'aya, naqi a-m-d-u-x-a.*
'No, not thus you will do, make them'[26]

23. Glosses and translations are given by Boas.

24. Note that Lower Chinook, like Upper Chehalis, is a VSO language. Hence, the order in negative sentences, with both a subject and an object NP, is expected to be NegVSO.

25. The verb *aq'Lāx* can be analyzed as follows: *a-* 'aorist'; *q'-* 'subject SOME ONE'; *-L-* 'object IT'; *-ā-* 'directive'; *-x* 'stem: TO DO'.

26. This gloss is given by Dyk (1933).

2.2 Upper Chehalis

As in Lower Chinook, in Upper Chehalis, the negative particle *míłta* precedes the verb and "it occurs as the first element in a sentence or embedded clause" (Kinkade 1976: 19), see (25a). There is some uncertainty with regard to the categorial status of the negative marker. Kinkade (1963a: 345) classifies *míłta* among particles, the "only major morpheme class which does not have affixes." On the other hand, Thomason (1983) and Kinkade (p.c., 1999) note that in all Salish languages, and in Upper Chehalis specifically, the negative marker is an intransitive verb with restricted inflectional properties. It is usually followed by a particular construction, a kind of nominalization, introduced by a prefix *s-* and optionally preceded by the indefinite particle *t*.[27] Based on examples found in Kinkade (1976), it can be concluded that the negator also precedes any free modal or tense markers that may occur in front of the verb, see (25b). On the other hand, the conditional particle *ʔáma?* 'if' precedes the negative marker in (25c).

(25) a. *Míłta t n-s-ʔəx-ci.*
not indef. my-cont.-see-you
'I don't see you'

b. *Míłta q'ał s-yucá-y-tt.*
not MOD IM-kill-TR-PASS
'He could not be killed'

c. *ʔáma? míłta t s-wi-ns ʔac-kʷəna-xʷ.*
if not indef. cont.-be-his stative-take-it
'If he doesn't take it…'

Kinkade (1976) provides a couple of examples with negative quantifiers, see (26) below. In both cases these have, like those in Lower Chinook and CJ, a transparent structure and consist of interrogative pronouns preceded by the negative marker *míłta*.

(26) a. *Míłta ʔé·nəm t qał s-kʷáxʷ-s.*
not how indef. subj. cont.-arrive-his
'There's no way to get there'

b. *Míłta pən-čá λa s-qʷət'-s t šəwl.*
not time-where unrealized cont.-burn-its indef. trail
'A trail will never burn'

27. The subject is possessive in these constructions which are always continuative and can be transitive. See Kinkade (1963a, 1963b, 1976) for more details on the complexities of the constructions following the negative marker and on two alternative negation strategies less common than the one presented here. See also Boas (1934).

2.3 CJ negation in comparison to negation in source languages

As shown in the previous sections, CJ negation shows basic syntactic similarity with Upper Chinook and Upper Chehalis, two of its source languages, in terms of the positioning of the negative operator and the structure of negative words. In this section, further illustrative examples are provided from CJ, Lower Chinook (L. Chi), and Upper Chehalis (U.Che), for several different constructions.

Examples (27) to (31) illustrate the positioning of the negative marker with respect to other clausal constituents. Further similarities between CJ and the two source languages under consideration, in particular Lower Chinook, are demonstrated: (27) shows that the question particle follows the negative marker in both CJ and L. Chi;[28] (28) shows that a WH-word precedes the negative marker in both CJ and L. Chi;[29] (29) shows that the modal particle precedes the negative marker in both CJ and L. Chi;[30] (30) shows that the conditional conjunction precedes the negative marker in L. Chi, U.Che and CJ; finally, (31) shows that the negative marker precedes the verb in imperatives in L. Chi, U.Che and CJ.

(27) Yes-no questions and negation

a. L.Chi *Nēkct na tnē'txix?*
NEG Qpcl know-it
'Do I not know it?' (Boas 1910/1911: 650)

b. U.Che *Nəwena we syóscex^{w}o?*[31]
'Are you (SG) working yet?' (Kinkade 1964: 59)[32]

---

28. According to Kinkade (1964: 59) *-na* is an interrogative suffix attaching to various syntactic constituents including particles "when no interrogative word [...] is present". Hence, it is "not strictly a verbal suffix" and "[it] goes on the word or words about which the question is being asked". The interrogative marker in Upper Chehalis does not seem to be necessarily in a "word second" position, as in CJ and Lower Chinook (according to examples provided in Boas 1910/11).

29. Kinkade (p.c. 1999) reports that the negative marker is never preceded by a WH-word in Upper Chehalis.

30. Upper Chehalis seems to be different in this respect. The modal marker normally follows the negative marker, as in the following example:

Míłta q'ał s-yucá-y-tt.
not MOD IM-kill-TR-PASS
'He could not be killed.'

31. There is no gloss for this example in Kinkade (1964).

32. Kinkade (1964) doesn't provide any examples where the interrogative marker *-na* follows the negative marker. In personal communication (1999), Kinkade confirmed the lack of such examples in the Upper Chehalis sources. This example is given here to illustrate the use of *-na* (see also Note 28).

c. CJ *Wek na msaika komtaks naika?*
NEG Qpcl 2PL know 1SG
'Don't you know me?'

(28) Wh-words and negation

a. L.Chi *Qā'daqa nēkct ā'nqatē amiō'tXam?*
why not long.ago you.told.me
'Why did you not tell me long ago?' (Boas 1894: 67)

b. CJ *Klaksta wek mitlait kopa Noe yaka ayas knim, . . .*[33]
who NEG stay PREP Noah 3SG big canoe
'Whoever didn't stay in Noah's big boat,. . . '

(29) Modal particles and negation

a. L.Chi *. . . pōs nēkct ē'ka atci'lxax.*
Mpcl not thus he-us-direct.-to.do
'. . . he would have not done so to us' (Boas 1910/11: 650)

b. CJ *Tlus wek msaika krai kopa naika.*
good NEG 2PL cry PREP 1PL
'You shouldn't cry for me'

(30) Conditional conjunction and negation

a. L.Chi *Qē nēkct maikxa imē'q!atxala, pōs nēkct ē'ka atci'lxax.*
if not your badness Mpcl not thus he.did.to.us
'If it had not been for your badness, he would not have done so to us' (Boas 1910: 650)

b. U.Che *?áma? míł ta t s-wi-ns ?ac-k^{w}əna-x^{w}.*
if not indef. cont.-be-his stative-take-it
'If he doesn't take it. . . ' (Kinkade 1976: 20)

c. CJ *Pus wek msaika eskom ukuk naika wawa. . .*
if NEG 2PL take DEM 1SG word
'If you do not accept my words,. . . '

(31) Imperatives

a. L.Chi *Nēkct mō'ya iau'a!*
not go there
'Don't go there!'

b. U.Che *Míłta t?a-s-yá!*
not indef.-your-cont.-go.home
'Don't go home!'

c. CJ *Wek eskom ukuk!*
NEG take this
'Don't take this!'

33. An example with a WH-question is not available, and I give this relative sentence (the relative pronoun is identical in form to the interrogative pronoun) as the closest correlate.

The facts presented in (27) through (31) suggest that the structure of CP in Chinookan and Upper Chehalis might be quite similar to that of CJ (see (19) and (21) above). Further examples are provided in (32)[34] to show that negative indefinites are formed by combining the negative marker with the interrogative pronoun in all three languages.

(32) Negative quantifiers in L.Chi, U.Che and CJ

a.	L.Chi	*nikct*	*ē'kta*	'nothing'
		NEG	what	
		nikct	*qa'nsix*	'not (any)how'
		NEG	(any)how	
b.	U.Che	*míłta*	*wa:*	'nobody'
		NEG	who	
		míłta	*tam*	'nothing'
		NEG	what	
		míłta	*pən-čá*	'never'
		NEG	when	
c.	CJ	*elo*	*ikta*	'nothing'
		NEG	what	
		elo	*klaksta*	'nobody'
		NEG	who	
		wek	*kansih*	'never'
		NEG	when	

In addition, as in CJ, only one negative operator per sentence is allowed in Lower Chinook and Upper Chehalis. Hence, the source languages, like CJ, do not allow negative concord, as illustrated in examples in (33) below. These examples also show that, in Lower Chinook and Upper Chehalis, as in CJ, the object negative operator in (33a) and adverbial negative operator in (33b) have to be preposed.[35]

(33) Negative quantifier position

a. L.Chi …*qē'wa* <u>*nikct*</u> <u>*ē'kta*</u> *L!ap* *agā'yax.*
…as not anything find I.did.it
'…because I did not find anything' (Boas 1894: 75)

b. U.Che <u>*Míłta*</u> <u>*ʔé·nəm*</u> *t* *qał* *s-k^{w}áx^{w}-s.*
not how indef. subj. cont.-arrive-his
'There's no way to get there'

c. CJ <u>*Elo*</u> <u>*ikta*</u> *masachi* *naika* *mamuk* […].
NEG what bad 1PL do
'I did nothing wrong…'

34. None of the sources give full paradigms of these items.

35. In all three languages objects normally follow verbs. The issue of their exact placement (i.e., whether they are VP internal or not) will not be addressed here.

Note that CJ and seemingly, Lower Chinook and Upper Chehalis, have properties corresponding to those identified by Zanuttini for a number of Romance languages (1997: 15ff.). Namely, Zanuttini has established that a) "pre-verbal negative markers that can negate a clause alone" cannot be used in front of a morphologically distinct imperative form; b) they always precede the pre-verbal pronominal clitic; c) they obligatorily co-occur with the negative indefinite following the verb; d) if the negative indefinite is preposed, the negative marker is not necessary.

CJ seems to have the last two properties.[36] Furthermore, the negative marker precedes the subject pronouns (which are, however, not clitics in the dialect of CJ under consideration). Finally, the negative marker precedes the imperative in CJ (which has no distinct imperative morphology). These properties also seem to be shared by Lower Chinook and Upper Chehalis, suggesting that the analysis presented here might be on the right track.

Following the preceding discussion, it can be concluded that CJ has considerable syntactic similarities with at least two of its most important Native American source languages: a) the syntactic position of the negative marker and its ordering in relation to other clausal constituents, b) the lack of negative concord, c) the positioning of negative indefinite NPs, and d) the inner structure of negative words.

3. Summary and discussion

The sentential negative marker *wek* in CJ is base-generated in the NegP within the expanded CP. CJ has no negative concord and the negative marker *wek* is a phrasal element in [Spec, NegP]. Negative quantifiers in CJ must raise to NegP overtly for feature checking. CJ negation was compared to the properties of negative constructions in Lower Chinook and Upper Chehalis, to the extent that data on these languages are available. It was concluded that sentential negation in CJ and in two of its important source languages have the following properties in common: a) negative marker precedes the IP, b) there is no negative concord, c) negative quantifiers are preposed, and d) negative quantifiers are composed of the negative marker and the WH-word.

The structural congruence of the dominant contact languages seems to have enabled the retention of the syntactic features characterizing sentential negation in CJ, in accordance with the proposals about the role of linguistic homogeneity in contact-induced language change made by Thomason and Kaufman (1988) and Singler (1988), among others. It is interesting that, unlike negation, the syntactic feature of word order wasn't retained in CJ, which is an SVO language, despite the fact that VSO was the dominant word order among the contact languages, as discussed in Vrzić (1997, 1999).

36. In contrast to the languages discussed by Zanuttini (1997), the preposing of negative indefinites in CJ, L.Chi and U.Che seems to be obligatory.

This difference in the retention of source language syntactic features in CJ suggests an interesting direction toward a further formulation of universal syntactic constraints on language contact and substrate influence in pidgin/creole genesis proposed in Vrzić (1999). Several linguistic factors could have contributed to the retention of source language features in the area of negation, unlike that of word order. First, the negative marker is a free morpheme in the contact languages considered in this paper. Second, while the negative marker is a functional morpheme associated with a functional category, this category contains interpretable features, i.e., those features that are relevant to the semantic interpretation of the sentence. Finally, the negative marker is base-generated in a structural position into which it would need to move in the course of syntactic derivation for feature checking, if it were not already in it, making the base-generation an economical structural property. None of these properties are shared by syntactic categories and processes governing basic word order in contact languages.

References

Boas, F. (1894). *Chinook Texts.* Bureau of American Ethnology, Bulletin No.20.

Boas, F. (1910/11). *Chinook.* Extract from Handbook of American Indian Languages (Bulletin 40), Part 1, of Bureau of American Ethnology. Washington DC: Government Printing Office.

Boas, F. (1934). A Chehalis text. *International Journal of American Linguistics, 8*, 103–110.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The View from Building 20: Essays in linguistics in honor of Sylvain Bromberger* (pp. 1–52). Cambridge MA: The MIT Press.

Culicover, P. (1991). Topicalization, inversion, and complementizers in English. In D. Delfitto, M. Everaert, A. Evers, & F. Stuurman (Eds.), *Going Romance and Beyond* (pp. 1–43). Utrecht: University of Utrecht.

Culicover, P. (1993). Evidence against ECP accounts of the *that*-t effect. *Linguistic Inquiry, 24*, 557–561.

Degraff, M. (1993). A riddle of negation in Haitian. *Probus, 5*, 63–93.

Dobrovie-Sorin, C. (1994). *The Syntax of Romanian: Comparative studies in Romance.* Berlin: Mouton de Gruyter.

Dyk, W. (1933). A Grammar of Wishram. Ph.D dissertation, Yale University.

Haegeman, L. (1995). *The Syntax of Negation.* Cambridge: CUP.

Haegeman, L. & Zanuttini, R. (1991). Negative heads and the Neg-Criterion. *The Linguistic Review, 8*, 233–251.

Haegeman, L. & Zanuttini, R. (1996). Negative concord in West Flemish. In A. Belletti & L. Rizzi (Eds.), *Parameters and Functional Heads: Essays in comparative syntax* (pp. 117–179). Oxford: OUP.

Hajda, Y. P., Zenk, H., & Boyd, R. (1988). The early historiography of Chinook Jargon. Paper presented at the 87th annual meeting of the American Anthropological Association, Phoenix, Arizona.

Jacobs, M. (1936). Texts in Chinook Jargon. *University of Washington Publications in Anthropology, 7*, 1–27.

Kamloops Wawa (1891–1923). *Nos. 1-507*. Kamloops BC: St. Louis Mission.

Kayne, R. (1998). Overt vs. covert movement. *Syntax, 1*(2), 128–191.
Kinkade, M. D. (1963a). Phonology and morphology of Upper Chehalis: I. *International Journal of American Linguistics, 29*(3), 181–195.
Kinkade, M. D. (1963b). Phonology and morphology of Upper Chehalis: II. *International Journal of American Linguistics, 29*(4), 345–356.
Kinkade, M. D. (1964). Phonology and morphology of Upper Chehalis: III. *International Journal of American Linguistics, 30*(1), 32–61.
Kinkade, M. D. (1976). The copula and negatives in Inland Olympic Salish. *International Journal of American Linguistics, 42*(1), 17–23.
Laka, I. (1990). Negation in syntax: On the nature of functional categories and projections. Ph.D dissertation, Massachusetts Institute of Technology, Cambridge MA.
Platzack, C. (1986). COMP, INFL, and Germanic word order. In L. Hellan & K. K. Christensen (Eds.), *Topics in Scandinavian Syntax* (pp. 185–234). Dordrecht: Reidel.
Pollock, J.-Y. (1989). Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry, 20*, 365–424.
Rivero, M.-L. (1994). Clause structure and V-movement in the languages of the Balkans. *Natural languages and linguistic theory, 12*(1), 63–120.
Rizzi, L. (1996 [1991]). Residual verb second and the Wh-criterion. In A. Belletti & L. Rizzi (Eds.), *Parameters and Functional Heads: Essays in comparative syntax* (pp. 63–90). Oxford: OUP.
Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of Grammar* (pp. 281–339). Dordrecht: Kluwer.
Rudin, C. (1985a). *Da* clauses, finiteness, and opacity. *Folia Slavica, 7*, 435–451.
Rudin, C. (1985b). *Aspects of Bulgarian Syntax: Complementizers and wh-constructions*. Columbus OH: Slavica.
Singler, J. V. (1988). The homogeneity of the substrate as a factor in pidgin/creole genesis. *Language, 64*, 27–51.
Thomason, S. G. (1983). Chinook Jargon in areal and historical context. *Language, 59*, 820–870.
Thomason, S. G. & Kaufman, T. (1988). *Language Contact, Creolization and Genetic Linguistics*. Los Angeles CA: University of California Press.
Vrzić, Z. (1996). The categorial status of modal *da* in Serbo-Croatian. In J. Toman (Ed.), *Annual workshop on formal approaches to Slavic linguistics* (pp. 291–313). Ann Arbor MI: Michigan Slavic Publications.
Vrzić, Z. (1997). A minimalist approach to word order in Chinook Jargon and the theory of creole genesis. In B. Bruening (Ed.), *MIT working papers in linguistics* 31: *Proceedings of the Eighth Student Conference Student Conference in Linguistics* (pp. 267–278). Cambridge MA: MIT, Department of Linguistics and Philosophy, MITWPL.
Vrzić, Z. (1999). Negotiating Features in Ppidgin/creole Genesis: Universals and contact influence in Chinook Jargon syntax. Ph.D dissertation, New York University.
Vrzić, Z. (forthcoming). *Chinook Jargon Texts and Grammar.* Munich: Lincom.
Zanuttini, R. (1991). *Syntactic Properties of Sentential Negation: A comparative study of Romance languages.* Ph.D dissertation. Philadelphia: University of Pennsylvania.
Zanuttini, R. (1994). Re-examining negative clauses. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, & R. Zanuttini (Eds.), *Paths Toward Universal Grammar: Studies in honor of Richard S. Kayne* (pp. 427–451). Washington DC: Georgetown University Press.
Zanuttini, R. (1997). *Negation and Clause Structure: A comparative study of Romance languages.* Oxford: OUP.

Zenk, H. (1984). Chinook Jargon and Native Cultural Persistence in the Grand Ronde Indian Community, 1856–1907: A special case of creolization. Ph.D dissertation, University of Oregon.

Sri Lankan Malay morphosyntax

Lankan or Malay?[1]

Peter Slomanson
City University of New York Graduate Center, USA

In this paper the convergence of Sri Lankan Malay on the grammars of Muslim Tamil and colloquial Sinhala is discussed. This convergence is incomplete and its extent is asymmetrical across syntactic domains. The verbal domain remains closer to the syntax of other contact Malay varieties, as demonstrated by the pre-verbal surface position of major functional elements, including temporal markers, modal elements, and negation. Greater convergence in the nominal domain follows a shift from pre-nominal to post-nominal adpositions. The conservative distribution of functional morphology in the verbal domain (predominantly pre-verbal) is associated with weak verb movement, contra Tamil and Sinhala. This reflects the verbal syntax of contact Malay varieties from Indonesia, from which most lexical material in Sri Lankan Malay has been retained.

---

1. This is a revised version of a paper I presented at the 7th International Symposium on Malay/Indonesian Linguistics in Nijmegen, the Netherlands in June 2003, and at the Society for Pidgin & Creole Linguistics meeting at the University of Hawai'i at Mānoa in August 2003. I owe a substantial debt of gratitude to Mohamed Jaffar, Hinaya Jainoor, and several other Colombo Malay informants, too many to name, who provided native SLM grammaticality judgments, to Burhanudeen Jalaldeen, Mohamed Tawfiq Mohamed Rihan and many other native speakers in Kirinda and Hambantota for confirming the syntactic and morphosyntactic validity of my examples in southeastern SLM, to Wasantha Amarakeerthi Liyanage for native Sinhala grammaticality judgments, and to Anne Bichsel-Stettler for her thesis and invaluable recordings of native SLM speakers in Sri Lanka. I would also like to thank Sander Adelaar, Corinna Anderson, Marcel den Dikken, Scott Paauw, and Ian Smith for feedback and discussion, Umberto Ansaldo for his manuscript, and Ana Deumert, Stephanie Durrleman, Theresa McGarry, John Singler, and anonymous reviewers for their editorial comments.

1. Introduction

The lexical inventory of Sri Lankan Malay (SLM),[2] both open and closed-class elements, remains largely Malay, but the language's grammar has come under substantial influence from co-territorial south Asian languages, which are typologically distant from the Malay varieties spoken in southeast Asia. This has led to the claim that the syntax of SLM is derived from the syntax of Lankan languages – particularly Tamil[3] – and its functional morphology is Tamil in all but its phonological shapes.[4] I will take the divergent position that the morphosyntax of the verbal domain in SLM is relatively non-convergent with the morphosyntax of the verbal domain in Tamil and Sinhala, compared with what we find for the nominal domain.[5] I will motivate this claim in part by showing that functional elements which modify SLM verbs are not just conservative (i.e. Malay) in form, but also unlike Tamil and Sinhala in their distribution. I will also claim that two properties which make the verb syntax of SLM appear as typologically Lankan as it does, OV surface orders and non-finite clauses, are part of the nominal syntax of the language. The nominal domain converges to a significantly greater extent with the grammars of Tamil and Sinhala than does the verbal domain.

2. Non-convergent predicate orders

In most of the literature on SLM, the data presented have included OV sentences exclusively.[6] Based on those data, it has been concluded incorrectly that SLM uniformly

---

2. Since first presenting this work, I have recorded native SLM from several communities, and have been struck by the degree of cross-dialectal unity. An important source of variation is the relative degree of Sinhala influence, which is greater in southeastern Sri Lanka than elsewhere, and greatest among younger speakers generally. Cross-lectally contrasting degrees of specifically Sinhala influence are to a large extent – though by no means exclusively – phonological and lexical, phenomena which are beyond the scope of this paper.

3. The variety of Tamil of relevance to the creation of Sri Lankan Malay is Sri Lankan Muslim Tamil, which I will refer to in my examples simply as Muslim Tamil. Historically the English glottonym has been Moorish Tamil, but the term Moorish is currently regarded unfavorably by speakers of the dialect. The glottonym in the dialect itself is [ʃo:nəm] (M. Jaffar, p.c.).

4. For example, according to Hussainmiya (1986: 22), "The contact situation between Sri Lankan Moor Tamil and SLM resulted not only in lexical borrowings, but it has pervaded all aspects of the latter's grammatical system."

5. All references to Sinhala are to the vernacular, as written Sinhala differs considerably and was not widely used in the SLM community until the second half of the twentieth century.

6. Exceptions include Saldin (1996) and one or two sentences in Bichsel-Stettler (1989).

manifests OV surface ordering in all syntactic contexts.[7] There is considerable variation in the relative ordering of verbs and their complements however. Across regional varieties, matrix verbs frequently appear to the left of nominal complements in pragmatically unmarked declarative contexts, creating root VO clauses, as demonstrated in (1a) and (2).[8]

(1) a. *itu babi su*[9] *gigit kičil anak naŋ.* SLM (VO order)
that pig PAST bite small child ACC-EXPER-DEF
'That pig bit the small child'

b. *babi-tu su gigit ana kačil.* AMBONESE MALAY[10]
pig that ASP bite child small
'That pig bit the small child'

The linear ordering contrasts with southeast Asian Malay (1b) which we find in these sentences are limited to *nominal* syntax. In (1a), the adjective is pre-nominal, rather than post-nominal as it is in Ambonese Malay and other non-Lankan varieties. This is attributed to influence from Tamil and Sinhala (Adelaar 1991:25). The determiner (*itu*) is also pre-nominal, as are Tamil and Sinhala demonstratives, although Ambonese

7. Thus, Hussainmiya (1986) writes: "The word order in Sri Lankan Malay follows exactly the Sri Lankan Moor Tamil. In non-copulative sentences it follows SOV structure." Similarly, Bakker (2000) states: "SLM is an 'SOV' language: in the basic sentence structure, the Subject comes first, the Verb comes last, and the Object is in between. Other Malay dialects have an 'SVO' construction."

8. Robuchon (2003) presents two examples of VO sentences in the speech of a young informant, introducing the syntactic variation displayed by his informant with the following comment.

> But one interesting feature of the speaker's behaviour was the use of a pattern-shift, making the norm double on certain aspects of the syntax: a norm resulted from the historical Tamil/Sinhala interferences, as we saw them; and another norm, that of a more standard Malay, which remained like buried in the depth of the dialect (and of the memory of the speaker) and re-emerges from time to time, not according to any complementary distribution of uses, but as a doublet composed of two syntactical patterns (a conservative one and a borrowed one) in competition for the same function.

9. The reduced form *su* of the aspect particle *sudah* is also present in other contact Malay varieties such as Ambonese Malay. The actual pronunciation in SLM is frequently [so] or [sə]; however, I have retained the conventional spelling *su*, which is found in all of the literature with the exception of Bichsel-Stettler (1989).

10. Ambonese Malay is a member of a group of Indonesian (Moluccan) contact varieties which Adelaar (1991) claims are plausible substrate candidates for SLM, based largely on phonological criteria. Influence from the Jakarta Malay of that period may actually have been greater than from Moluccan varieties (S. Paauw, p.c.). Morphosyntactic parallels between Jakartan and Moluccan Malay varieties justify the use of Ambonese examples though.

Malay demonstratives can also occur pre-nominally (van Minde 1997: 147),[11] so the non-Malay origin of this feature is harder to establish. The sentence in (2) is indistinguishable, at least in surface form, from what we find in Malay vernaculars spoken outside of Sri Lanka generally.

(2) *anak su makan*[12] *nasi.* SLM (VO order)
child PAST eat rice
'The child ate rice'

SLM appears in all other respects to constitute a language of the type denoted by the term *converted language*, used in Bakker (2000: 29–35) to categorize SLM and Sri Lankan Portuguese creole, including postpositions, prenominal adjectives, and prenominal relative clauses. However the VO sentence type, given the contrast between its VP head-complement order and the unmarked OV order in Muslim Tamil and Sinhala, causes SLM to fall short of converted language status. Bakker uses the term converted language to refer to a contact language displaying an inherited lexicon and a borrowed grammar.

Adelaar (1991: 30–31) and Bakker (2000: 29–35) report on data lacking VO orders. Those data led both authors to conclude that only OV orders are available in SLM. Even in their OV data, however, we find functional morphemes such as tense markers consistently appearing to the left of an associated main verb. Which reflects the surface position of negation and aspect relative to the verb in non-Lankan Malay varieties.

In a generative account, OV orders in SLM can be produced by raising object DPs[13] to the left of verbs. In VO clauses, the finite verb may appear to be raising to a higher functional position than in OV clauses, preventing object raising from yielding the OV surface strings for finite clauses; however object raising alone can account for the variation without recourse to variable verb movement.[14] The syntax of temporal elements and negation in relation to the verb never varies cross-dialectally in SLM based on object position. In a syntax-driven account of morpheme order, this

11. According to van Minde, pre-nominal demonstratives are far less frequent however, and unlike the post-nominal variant, cannot be 'strongly' stressed.

12. Word-final nasals, normally velar, tend to develarize before [n] in rapid speech.

13. I adhere to the convention of using the term DP to refer to nominal structures generally, in accordance with the view in certain recent syntactic frameworks that a noun phrase is the complement of its determiner (as a clause is the complement of its complementizer). For the purposes of this paper, DP can be read as NP (or noun phrase) without rendering the description inaccurate.

14. Slomanson (2001) parenthetically analyzed VO/OV variation in SLM root clauses as an example of variable object movement resulting from grammatical change under language contact conditions.

fact should rule out a role for verb movement in explaining verb-complement order variation.[15]

The variable position of nominal objects suggests on-going competition between a residual grammar featuring categorical VO ordering and a more recent grammar which favors OV orders. The validity of this grammatical competition approach is strengthened by the lack of pragmatic contrast between the two orders and by an on-going shift in the frequency of the two orders in favor of the OV variant. The fact that the frequency of OV root orders has increased over time (i.e. across generations) is evidence that root VO/OV variation is unstable in SLM. Typological changes involving a range of phenomena including head-complement orders tend not to occur simultaneously across a grammar. Certain kinds of variation are easily supported by a single grammar. For example, we find both closed class modal prefix of ability (*bolɛ*) on SLM verbs and a modal adjective with identical meaning and identical phonological shape.[16] These do not however constitute typologically conflicting options, which would render the resulting variation unstable. OV/VO variation, by contrast is more likely to be unstable over time, particularly because pragmatically-unmarked VO orders are typologically discordant with the Dravidian-type grammar on which SLM has been converging.

3. Embedded clauses

Consider the embedded clause in the sentence in (3) from Saldin (2001:65):

(3) *emak* *sɛ-daŋ* *panggil* [***ruma*** ***mə-sapu*** ***naŋ***]. SLM
mother me-to call house NONFIN-sweep P
'Mother calls me to sweep the house'

We notice three facts about this clause which are reflective of south Asian SOV grammars, but which contrast starkly with southeast Asian Malay grammars.

---

15. However in a syntactic account in which there is an association between features and lexical items and corresponding functional heads, verb movement (i.e. of a verb stem) as a means of deriving complex verbs can motivate the position of the verb (stem) relative to its functional affixes.

16. The modal-marked verbal predicate construction in SLM:

(i) *Farida* *čingla* *nyanyi*-pədə *bolɛ-nyanyi.*
Farida Sinhala sing-PL MOD-sing
'Farida can sing Sinhala songs'

The modal predicate construction in SLM:

(ii) *Farida* *čingla* *nyanyi*-pədə *mə-nyanyi* *naŋ* *bolɛ.*
Farida Sinhala song-PL NONFIN-sing P MOD
'Farida can sing Sinhala songs'

- The surface ordering of word-level constituents within the embedded clause is of canonical OV type.
- The embedded verb is morphologically marked as non-finite.
- There is a complementizer-like element at the right edge of the construction (the adposition *naŋ*).

These facts indicate a significant degree of convergence on Muslim Tamil and Sinhala grammars. Note that the position of the non-finite clause itself (in example (3)) is unlike what we typically find in Muslim Tamil and Sinhala. Its position relative to the matrix clause is right-peripheral, rather than center-embedded. In OV languages such as Tamil, the unmarked option is for embedded clauses to occur to the left of a matrix clause, as in (4).

(4) *Farida* [***varu-v-aa***] *ʃɛnn-aa.* MUSLIM TAMIL
Farida come-FUT-3S-QUOT say-PAST-3S
'Farida said she he would come'

The infinitival Sinhala clause in (5) is also center-embedded.

(5) *gæ:nu* [***eloolu*** ***ga-nnə***] *ya-nəwa.* SINHALA
women vegetables buy-INF go-PRES
'Women are going to buy vegetables'

Center-embedding is optional in SLM. Hussainmiya (1986) treats *mə-naŋ*-marked SLM verbs simply as infinitives. In Section 4.3, I will propose a different analysis.

4. The verbal domain

4.1 Negation

As we can see from the Tamil example in (6) and the Sinhala example in (7), negation is right-peripheral in both languages (in matrix contexts in Sinhala[17]). In Tamil it is part of the complex morphology of the verb.

(6) *var-e-la.* MUSLIM TAMIL
come-MOD-NEG
'He cannot come'

(7) *eyaa mas ka-nne næ:.* SINHALA
3S meat eat-PRES NEG
'He doesn't eat meat'

By contrast, SLM negation binds to the immediate left of a verbal predicate, as in (8).

17. In Sinhala, negation is marked on the left edge of a lexical verb in certain non-root clauses.

(8) *Ali nyari ikaŋ tər-makaŋ.* SLM
Ali yesterday fish NEG-eat
'Ali didn't eat fish yesterday'

Negation takes precedence over other functional morphemes, such as tense markers, when there is competition for the pre-verbal position (as in (8)), since the overt expression of negation is restricted to this position, and must be expressed grammatically.

The negation element *tra* follows predicate adjectives.[18] Exceptions in which a negator precedes an adjective arise in a small number of lexical exceptions (cf. examples (14) and (15)), when an adjective is interpretable as a verb. When an adjective is interpretable as a noun, the negation element is *bukaŋ*. It is also possible for *tra* to appear immediately after a verb, but the interpretation is then markedly aspectual. This can be translated as 'not ever' (in the past).[19] The fact that ordinary verbal negation precedes the negated item, as in Malay dialects generally, although negation of an adjectival predicate usually follows the negated item, as it does in Tamil and Sinhala, is evidence of the relative syntactic conservatism of the verbal domain in SLM.[20]

SLM does not feature movable negation phrases containing a negation morpheme as their head. While the distribution of negative polarity items is relatively free, the presence of a negation head is obligatory. The negation morpheme must be prefixed to a verb or other predicate element.[21] In the case of open-class predicate adjectives,

---

18. The following example demonstrates this:

(i) *Farida gumuk tra.*
Farida fat NEG
'Farida is not fat.'

A comparative construction, calqued on Muslim Tamil and Sinhala, requires the use of a nominal negator with the adjective:

(ii) *Farida bukan gumuk, Rashida.*
Farida NEG fat Rashida
'It's not Farida who's fat, but Rashida who is.'

19. This is the negation of auxiliary *ada* which is a marker of perfect aspect. The auxiliary which must be expressed if the perfect verb occurs in a matrix clause. There is an analogous construction in Tamil and Sinhala, respectively; however in those languages the participle is not variably marked pre-verbally (i) as it is in SLM. The analogous existential auxiliaries are *iru* in Tamil and *tiyɛnəwa* in Sinhala.

(i) *Farida buk (as-) tulis ada.* SLM
Farida book (ASP-) write AUX
'Farida has written a book (at some time)'

20. More detailed discussion of the form and distribution of negation elements can be found in Slomanson (2005).

21. The following examples illustrate this generalization:

the negation marker *tra* is a contraction of *tər-ada* (neg-existential auxiliary verb), so although a negation head is required to appear at the right of the predicate adjective, the generalization holds (cf. footnote 18).

4.2 Tense and aspect

SLM verbs are pre-verbally tense-marked, as in (9) through (11).

(9) *pompaŋ kantor naŋ* ***su*** *pi si?* SLM
woman office to PAST go Q
'Has the woman gone to the office?'

(10) *pompaŋ* ***arə*** *makan nasi.* SLM (VO order)
woman PRES eat rice
'The woman is eating rice'

(11) *pompaŋ* ***anti*** *makan nasi.* SLM (VO order)
woman FUT eat rice
'The woman will eat rice'

There is no pre-verbal temporal morphology in Tamil, whose morphologically complex verbs encode a range of functional information, including tense, aspect, mood, agreement, and negation, by affixation to the *right* of a lexical or auxiliary verb rather than to the *left* of the verb stem. The fact that a functional element such as tense has not been reanalyzed as a *post*-verbal element in SLM, whether freely or affixally, demonstrates the syntactic non-convergence of the Malay verb system on the Tamil system, although the accretion of overt tense affixation[22] demonstrates functional-semantic convergence. Sinhala influence is not indicated, given the fact that future tense is only marked in the spoken variety when there is a first person agent, contra Tamil and SLM, and given the fact that Sinhala past tense is marked by *ablaut* and other phonological alternations, rather than simply by affixation.

(i) *Farida tər-tulis nyari atupon.*
Farida NEG-write today something
'Farida isn't writing anything today'

(ii) *Farida atupon nyari tər-tulis.*
Farida something today NEG-write
'Farida isn't writing anything today'

(iii) **Farida arə-tulis nyari tra atupon.*
Farida TNS-write today NEG anything
'Farida is not writing anything today'

22. I take these as descriptively 'affixal', since they are phonologically – dependent and restricted to pre-verbal position. I do not analyze them as inflections though, for theoretical reasons beyond the scope of this paper.

The element *ada*, one of the few TMA elements in SLM with no temporal analog in standard Malay – etymologically an existential verb – has been reanalyzed in SLM as a present tense marker, occurring in a number of variant phonetic shapes (*ada*, *arə*, and [r]). Ambonese Malay, presumably not in protracted contact with South Asian languages, features this functional element, but only as an aspect marker, as in (12).

(12) *De blong ada makang.* AMBONESE MALAY
3s not yet PROG eat
'He's not eating yet' (van Minde 1997: 231)

The analogous SLM sentence is distributionally similar,[23] although Ambonese Malay *ada* is (vacuously in this example) syntactically independent of *makaŋ*. In SLM, the tense marker *ada* (*arə*) cannot be separated from *makaŋ*.[24] *Post*-verbal *ada* is a marker of perfect aspect in SLM. Like the pre-verbal marker, it must be adjacent to the verb. Its post-verbal position suggests that the syntax of temporal elements in SLM has partially converged with that of the major Lankan languages, however incompletely (given the presence of *pre*-verbal temporal marking). The verb is variably preceded by a participial prefix (*abis*, *as*). Post-verbal *ada* is an auxiliary verb in SLM, which is analogous with *iru* in Tamil and *tiyɛnəwa* in Sinhala, both of which are existential in other contexts, as is preverbal *ada* in SLM.

Functional morphemes that mark tense on SLM verbs must be expressed (although certain speakers omit them occasionally) in unnegated verbal matrix clauses in which the verb is not otherwise marked by a modal morpheme. They obligatorily precede the finite verb,[25] although a number of the most common verbs may be variably unmarked. The pre-verbal surface distribution of functional morphemes in the verbal domain is typical of non-Lankan Malay varieties in which adverbial particles

23. The adverbial phrase *lai tra* for "not yet" is far more common in SLM than is *bloŋ*.

24. Pre-verbal *ada* is not in fact aspectual in SLM, having been reanalyzed as present tense; however, aspectual interpretation is implicit in this example, since intransitive *makaŋ* does not permit a punctual reading.

25. The one claim to the contrary of which I am aware comes from Robuchon (2003): "...the most largely used pattern for the past is that resorting to the particle *-le* suffixed to the verb." Robuchon's data come from a young Colombo speaker, possibly a semi-speaker. My own informants rate this construction as highly unlikely. The periphrastic perfect construction consists of a participle and an auxiliary verb taking the form ASP-V (NEG-)AUX (the auxiliary is *ada*). In rapid speech, this is sometimes produced as V AUX and the shape of the auxiliary reduced to *da*. This may create the impression that there is an actual past tense suffix, since the construction refers to events that have already taken place. If this is reanalyzed as a past tense marker, it will yield an incomplete paradigm, since there is no syntactically parallel construction which refers to present or future events or states.

(i) *Farida čingla nyanyi-atu as-nyanyi ada.*
Farida Sinhala song-DET ASP-sing EXPER
'Farida has sung a Sinhala song (before, at some point)'

cognate with SLM functional elements mark aspectual contrasts. Pre-verbal temporal markers are not found in Tamil or Sinhala; rather, temporal morphemes follow the verb, mostly as affixes. SLM tense elements such as *su* (past), *nya*(*ŋ*) (past), *e*(*ŋ*) (past, southeastern dialect), *arə* (present), and (*a*)*nti* (future),[26] are themselves affix-like, being obligatorily left-adjacent to a lexical verb.[27] This runs counter to what we find in non-Lankan Malay varieties. A verb prefixed with *su* is an inflected V. Note that in Ambonese Malay, which *also* features phonologically reduced TMA elements such as *su* (from *sudah*), no adjacency of *su* and verb is required. Ambonese Malay *su* occurs at the left edge of a functional complex which includes negation (van Minde 1997: 227–232). SLM does not permit the stacking of functional elements at the left (or right) edge of a finite (or non-finite) verb. It is not possible to separate SLM *su*, *nya*(*ŋ*), *e*(*ŋ*), *arə*, or (*a*)*nti* respectively from a verb. In casual prosody, these elements are never accented, and the vowel is typically reduced, so that in the case of *su*, the vowel nucleus is often reduced to schwa or deleted. The right edge of a finite SLM verb is frequently uninflected.

Paratactic sequences of verbs referring to sequences of events require repetition of the initial marker on the non-initial verb, even where this is redundant. Free-standing functional elements in INFL do not commonly exhibit this requirement cross-linguistically, whereas phonologically-dependent functional elements such as tense affixes typically do.

The syntax of full and phonologically reduced variants of the Malay perfect marker *abis* (*asa*, *as*, *ai*) depends on whether the element is expressed in a root or non-root clause. The non-root context, in which it is a participial element encoding perfect aspect, is exemplified in (13).

(13) *sɛ* [***məsigit naŋ as pi***] *ruma naŋ su dataŋ.* SLM
1s mosque to ASP go house to PAST come
'Having gone to the mosque, I came home'

As we find with *ada*, the marker can occur both pre-verbally and post-verbally. It is not reducible in post-verbal position.

4.3 Non-finite SLM clauses as nominalizations

The asymmetry between finite SLM clauses and non-finite (obligatorily OV) SLM clauses in the VO grammar may be attributable to a verbal and nominal domain contrast. Recall the claim that the verbal domain in SLM is more Malay-like than the

Further discussion of the syntax of *abis* as diagnostic of verb movement asymmetries across clause types can be found in Slomanson (2005).

26. This element is a temporal adverb in standard Malay/Indonesian with independent distribution. It has been grammaticalized as a tense marker in SLM, and cannot be used adverbially.

27. They are optional with the frequent verbs of motion *dataŋ* 'come' and *pi* 'go'.

nominal domain, given the surface distribution of tense markers and negation. The *mə-* and *-naŋ* elements phonologically dependent on verbs, taken by Hussainmiya as infinitival marking, do appear to instantiate a finiteness opposition unknown in non-Lankan Malay varieties of Malay;[28] however, there is reason to believe that verbs beaning these morpheme are not simply infinitives.

In phrases containing an apparently non-finite verb, the verb in question has been nominalized. The *mə-* in SLM *naŋ*-clauses is obligatory[29] for all verbs, *irrespective of argument structure*,[30] as in (14), in which the nominalized verb lacks an internal argument.

(14) [***Ali mə-dataŋ naŋ***] *tər-bai.* SLM
Ali NONFIN-come P NEG-good
'Ali's coming is not good'

The SLM construction may function as a subject, as in the intransitive example in (14), and as in the transitive example in (15). There is no comma intonation between the nominalized clause and the matrix predicate, and that position should not be interpreted as containing a null subject. Furthermore, the nominalization construction itself can be the complement of a postposition (e.g. *dupaŋ* 'before' and *blakaŋ* 'after').

(15) [***Ali sigrɛʈ mə-minuŋ naŋ***] *tər-bai.* SLM
Ali cigarette NONFIN-drink P NEG-good
'Ali smoking cigarettes is not good'

Sinhala is most likely to have provided the model for the construction in (14) and (15).[31] Nevertheless, it is worth considering possible parallels in Muslim Tamil. Muslim Tamil does productively nominalize clauses by the case-marking of verbs; however,

28. By contrast, non-Lankan varieties of Malay do not display overt morphological reflexes of finite status, such as bound tense markers which contrast with markers of non-finiteness. Sander Adelaar (p.c.) comments that there is ongoing debate among Austronesianists as to possible contrasting degrees of (morphologically-conditioned) temporality between verb constructions in Malay dialects in Indonesia and elsewhere, however the SLM finiteness contrast is perfectly explicit, as in Tamil and Sinhala, and can safely be attributed to contact with those languages.

29. Certain speakers, particularly younger ones, variably omit the *mə* morpheme. This is informally interpreted by my informants as evidence of semi-speaker status. I cannot say whether the variable deletion of this element should be treated as idiolectal or as evidence of a local change in progress (i.e. in a particular dialect).

30. Transitive *and intransitive* predicates in the SLM constructions containing the prefix *mə-* are both grammatical, contra standard Malay verbs prefixed by *me-*.

31. A follows from this that the construction is relatively recent, given the fact that strong influence from Sinhala probably began later than Muslim Tamil influence, and given the fact that the Sinhala construction in (18) is most strongly associated with Kandy, which was not politically integrated with the rest of the country until the nineteenth century. The diachrony is in need of further investigation, a difficult prospect given the lack of vernacular texts.

of the range of nominalization constructions in that language, there is no direct analog with the SLM *mə*-V-*naŋ* construction. The closest semantic and syntactic analog in Muslim Tamil is case-marked based on thematic status (i.e. +accusative) and inflected for tense.[32] The stem/affix order is V-tense-nominal-case. The tenseless gerundial construction in Tamil does not permit overt subjects.

The SLM *mə*-V-*naŋ* construction, by contrast, does not mark the thematic status of its maximal projection relative to a matrix predicate, and is allowed to contain an overt subject. In this sense, it is a compromise between the nominalization options we find in Tamil. In the non-finite type of sentential complementation in (14) and (15), the nominalizer *naŋ* functions as an assigner of unspecified case. OV orders are obligatory in these nominalizations. The clause is left-adjoined to *naŋ*.[33] Note that in Muslim Tamil (and in Tamil generally), infinitival clauses can occur as (negated) independent declarative clauses, as in (16):

(16) *avə ippo paɖam paa-kə poo-r-Illə* MUSLIM TAMIL
3SF now picture see-INFIN go-INFIN-NEG
'She does not go to a cinema now'

(17) **dɛ karaŋ baioskop naŋ jaŋ-pi naŋ.* SLM
1S now cinema P NEG.NONFIN-go P
'She does not go to a cinema now'

No such option exists for non-finite clauses in SLM (17), nor does the option exemplified in (16) exist in Sinhala.[34] In Sinhala, non-finite verbs can be dative-marked in nominalized embedded clauses, as in (18).

(18) *Ali* [***sindu-ak kiya-n-ʈə***] *a:wa:.* SINHALA
Ali song-INDEF say-INFIN-DAT came
'Ali came to sing a song'

(19) *Ali* [***nyanyi atu mə-nyanyi naŋ***] *su-dataŋ.* SLM
Ali song INDEF NONFIN-sing P PAST-come
'Ali came to sing a song'

Sinhala and SLM non-finite nominalizations correspond in permitting overt subjects.

---

32. According to Lehmann (1989: 300): "Both in terms of occurrence and assumption of syntactic functions, the three types of untensed verbal noun clauses are very much restricted in Modern Tamil. They do not occur in all NP positions, do not assume all syntactic functions of an NP, and do not occur with all case markers."

33. Hussainmiya (1986) treats both *mə-* and *naŋ* as infinitival affixes. It is reasonable to analyze *mə-* as an infinitival marker, appearing as it does in the same surface position relative to a verb as tense does. However *naŋ* – an adposition within ordinary PPs and an apparent case affix on nouns – is not an infinitive marker.

34. This does not indicate transfer in the sense of the actual calquing of constructions.

Sinhala also manifests a lexical constraint which is reflected in the SLM construction. The clausal complement of a perception verb cannot be nominalized with *-ʈə* (20a). The analogous SLM construction (21a) is similarly ungrammatical. Tense renders the embedded verb grammatical in both Sinhala (20b) and SLM (21b), but the case marker then changes from *naŋ* to *nya* in SLM, at least in the variety associated with the Colombo area.

(20) a. **mamə* [***Ali ya-n-ʈə***] *balə-nəwa.* SINHALA
1s Ali go-INFIN-DAT see-PRES
'I see Ali going'

b. *mamə* [***Ali ya-nəwa***] *balə-nəwa.* SINHALA
1s Ali go-PRES see-PRES
'I see Ali going'

(21) a. **sɛ* [***Ali mə-pi naŋ***] *liyat.* SLM
1s Ali NONFIN-go P see
'I see Ali going'

b. [***Ali arə-pi-nya***[35]] *sɛ arə-liyat.* SLM
Ali PRES-go-NOM 1s PRES-see
'I see Ali going'

SLM non-finite nominalizations are negated with *jaŋ* (southeastern *jɛŋ*). This is a reduced form of the element *jaŋaŋ*, which is associated in this and non-Lankan Malay varieties with negative imperatives. The elements *jaŋ* and *mə* cannot co-occur, as in (22).

(22) [***jaŋ*** (****mə-***)***taksir naŋ***] *Ali su pi.* SLM
NEG NONFIN-think P Ali PAST go
'Ali went without thinking'

Consider the SLM sentence in (23a) from Adelaar (1991:31):

(23) a. *de* [***attu pɔhɔŋ mɔ jadi-kiŋ***] *arə pi.* SLM
3s one plant will become-CAUS PRES go
'She's going to grow a plant'

Adelaar incorrectly glosses *mɔ*[36] as 'will', on analogy with the future tense marker in Ambonese Malay. Saldin (1996:65) refers to *mə-* as the SLM infinitive marker, reanalyzed from the standard Malay *meN-* prefix, which marks transitivity. That etymology

35. SLM, like Tamil and Sinhala, also features nominalized clauses, such as this one, which are tense-marked. The suffixes *naŋ* and *nya* do not have identical syntactic properties. Their functions in relation to nominalized clauses contrast, with *naŋ* marking non-finite clauses, and *nya* marking tensed clauses in Colombo Malay. There is considerable variation in the phonological shapes of these elements, however, with both forms – but particularly *naŋ* – sometimes reduced to *na*.

36. The form *mɔ* is a variant pronunciation of *mə*.

is unlikely. It is much more likely that SLM *mə-* is diachronically related to the modal element with which Adelaar associates it synchronically. In modern SLM, *mɔ* (mə-) in (23a) is not a future tense marker, but simply a marker of non-finite status. For evidence that this is the case, consider the modified example in (23b).

(23) b. *dɛ* [***pɔhɔŋ mɔ-jadi-kiŋ***] *su suka tadi.* SLM
3s plant NONFIN-become-CAUS PAST like then
'She liked to grow plants then'

The proposition in (23b), although referring to a past state, is grammatical. Note that *attu pɔhɔŋ mɔ jadi-kiŋ* occurs to the left of the semi-aspectual matrix verb *pi* in (23a), but the relative ordering of *pi* and the tense marker *arə* in the matrix verb cluster does not change. Only verbs with the *mə*/*mɔ*-prefix can raise to the left of other verbs and verb-modifying functional elements. The cluster *arə-pi* is semantically analogous with 'is going to'. The morpheme *mə* marks the absence of tense.

We have seen that SLM, with non-finite clauses, differs from other Malay varieties. These clauses are nominalizations, which are robustly attested in Tamil and Sinhala. However the morphosyntax of the SLM clauses imperfectly reflects that of analogous constructions in Muslim Tamil. Non-finite clauses in SLM are arguably associated with the nominal domain. As is the case with PP, to which the construction is diachronically related via reanalysis of the Javanese preposition *naŋ*, these clauses are rigidly head-final, taking *naŋ* to be the head of the nominalized clause. Note that a nominalized *naŋ*-clause can itself be the complement of an adposition (postposition), such as *blakaŋ* 'after'.

5. The nominal domain

From a typological perspective, the presence of postpositions and pre-nominal adjectives is suggestive of an SOV grammar in line with Muslim Tamil and colloquial Sinhala,[37] as demonstrated by the SLM PP in (24).

(24) *itu nakal anak sama* SLM
DET naughty child with
'with that naughty child'

5.1 Case markers as inflection?

In SLM, as in Tamil, accusative case-marking of non-verbal object DPs generally correlates with definiteness, as in the SLM examples in (25a) and (25b), although accusative case markers may also appear on nouns that can be interpreted as indefinite. The

37. SLM does not feature the obligatory indefiniteness marking that characterizes Sinhala nouns however.

sentence in (25b) refers to particular medicine whose identity is pre-established in discourse. An accusative/dative-marking split in SLM – paralleled in Tamil – is in evidence across dialects, although the phonological shape of the morphemes varies. Accusative case is not consistently marked in Sinhala.

(25) a. *Ali obat makaŋ.* SLM
Ali medicine eat
'Ali takes medicine'

b. *Ali obat-nya*[38] *makaŋ.* SLM
Ali medicine-ACC/DEF eat
'Ali takes the medicine'

Whereas the evidence for the affixal status of SLM postpositions is based entirely on adjacency facts, case inflection in Tamil triggers morphophonological changes to a case-bearing noun, as in (26).

(26) *viiɖu + ukku → viiʈʈ-ukku* TAMIL
house DAT house-to

The post-possessor possessive function of the morpheme *pɛ* has been universally acknowledged in the existing literature on SLM. It is important to note that syntactically and phonologically analogous forms are present in Moluccan and other Malay varieties, so this is not a Lankan contact innovation although Tamil and Sinhala happen to feature post-possessor possessive marking as well.

5.2 Functional reanalysis: Post-nominal *yaŋ/nya(ŋ)* as a case marker

I assume that the function of *yaŋ* as an accusative case marker in several SLM dialects originated in its use in WH-questions involving the extraction of patient objects in Jakarta Malay. The use of *yaŋ* in Jakarta Malay relative clauses, involving the extraction of subjects, is not implicated in the reanalysis of *yaŋ* as an accusative case marker.[39] In Jakarta Malay, which strongly influenced the creation of SLM, *yaŋ* is used

38. The accusative case suffix *-nya* can be realized in Kirinda Malay not just as *-yaŋ*, but as *-æŋ* (B. Jalaldeen, p.c.).

39. The anthropologist Bichsel-Stettler (1989) glosses the numerous tokens of *nya* and *yaŋ* (*yeŋ*) found in her corpus simply as linkers, as she was uncertain of their function (A. Bichsel-Stettler, p.c.).

Ansaldo (2005: 24) correctly analyzes *yaŋ* as a marker of accusative direct objects (i.e. patients) in Kirinda Malay; however he incorrectly claims that Kirinda Malay is the only variety that features the accusative/dative split which he treats as an isolated innovation. As I have not yet found a speaker in any of the existing varieties who does not have the split, Ansaldo's claim is unsupported. In an effort to show that other SLM varieties do not feature the split, Ansaldo relies on Smith, Paauw, and Hussainmiya (2004) who do not distinguish between accusative and dative case markers in SLM. However their data *also* come from Kirinda, a fact which Ansaldo himself

obligatorily in questions involving objects because the varieties in question lacked (and still lack) syntactic WH-movement. The use of *yaŋ* permits the extraction of (passivized) objects. I assume that *yaŋ* was therefore reinterpreted as a generalized accusative marker in declarative contexts. This is clearly due to influence from Tamil, presumably the result of L2 acquisition of Malay by native Muslim Tamil speakers, rather than due to influence from Sinhala. Accusative marking is used consistently in SLM as in Tamil, whereas accusative marking in Sinhala is optional. The WH-movement-based diachronic hypothesis proposed here establishes a logical mapping between the function of *yaŋ* in antecedent Malay varieties and the function of *yaŋ* in SLM. Without this connection, it is unclear why SLM would select *yaŋ* from the set of Malay phonological shapes available to its speakers.

The phonological shape of the *yaŋ* element in Jakarta Malay has varied historically and still does. Significantly it often begins with [n]. The SLM accusative marker assumes the forms *yaŋ, nya,* and *nyaŋ.* The variation in the shape of the Jakarta Malay relativizer accounts for the variation in the shape of the SLM accusative marker. These are historical allomorphs, although their use currently varies across regional SLM varieties, rather than being used variably by all speakers. The fact that the determiner *nya* encodes definiteness in Indonesian Malay varieties and that the accusative marker *nya* encodes definiteness in SLM creates the impression that *yaŋ* and *nya* are etymologically unrelated, since *yan* is a relativizer in Jakarta Malay and therefore functionally differentiated etymologically. It is more plausible however that the SLM *nya* is not the Malay definite determiner (*nya* cannot be suffixed to subjects in SLM), but rather that it is historically an allomorph of *nyaŋ*, in which the nasal coda can be deleted. The argument based on WH-movement is supported by the fact that WH-fronting of patient noun phrases in SLM *still* requires the use of *yaŋ*, which can alternate with the form *nya*,

---

reports elsewhere in his paper. While it is possible that certain case markers were misglossed in Smith, Paauw, and Hussainmiya's corpus (S. Paauw, p.c.), it is unclear why Ansaldo uses fairly recent data from the small village of Kirinda as a basis for comparison with recent data from the same village. Consider the following statements from Ansaldo (2005):

> The accusative marker thus appears to be an innovation of the variety of Kirinda… (Ansaldo 2005: 15)

> Table 3 presents the nominal marking of SLM adapted from Smith et al. (2004: 11). Interestingly, their account is largely based on data collected in Kirinda in the 1970s (I. Smith, p.c.) and therefore offers excellent comparative material. (Ansaldo 2005: 18)

> The conflation of Dative-Accusative observed in SLM has split into two different forms in Kirinda Java, paralleling the adstrate typology. (Ansaldo 2005: 19)

> In using different markers for Accusative and Dative, Kirinda Java has developed one step further than other SLM varieties that are reported to conflate these two functions (Smith et al. 2004). (Ansaldo 2005: 24)

The SLM of older speakers shows us that the accusative/dative split is not a recent development in Kirinda, and that the split was and remains ordinary usage in other dialect areas as well.

now interpretable in this syntactic context as a phonologically variable accusative case marker.[40] Further diachronic evidence for the course of events suggested by this analysis lies in the fact that SLM, like Jakarta Malay, still uses an overt relativizer, prefixed to the verb (31). The phonological shape of the SLM relativizer is not *yaŋ*, but *nya*.

5.3 Functional extension: Post-nominal *naŋ* as a case marker

The element *naŋ*,[41] both etymologically and persistently an adposition, is also a case marker of definite nominal objects, of experiencer subjects, and is a case marker of nouns in complex PPs. It most frequently appears as an ordinary dative marker, but it also marks the experiencer complement of a number of verbs, as in (27).

(27) *polis labrak maliŋ naŋ.* SLM (VO order)
police assault thief DAT.EXPER.DEF
'The police assault the thief' (Saldin 2001: 59)

A *semantic* relationship between the Tamil dative case inflection *-ukku* or the Sinhala *-ɭə* and SLM *naŋ* is plausible, because the dative case inflection is used in Tamil and Sinhala to indicate movement toward a destination. This corresponds with the role of *naŋ* as an allative postposition, which corresponds with its function as an allative preposition in Javanese, its most likely source. The Tamil dative inflection is also used in the existential possessive construction in (28), and the Sinhala example in (29), calqued in SLM, as we can see in (30).

(28) *en-akku ʃali iki.* MUSLIM TAMIL
1S-DAT money EXIST
'I have money'

40. This is as in the following SLM example, in which case-marking of the WH-word is obligatory:

(i) *apa-yaŋ/nya loraŋ arə-makaŋ?*
what-ACC you PRES-eat
'What are you eating?'

The fact that the object now has and requires case morphology and the question whether syntactic WH-movement is still disallowed, resulting in a construction which should be analyzed as a WH-cleft, are separate matters. I will not pursue the second one further in this paper, but I take it to be an important matter for future research. Note also that productive passive verbal morphology (Jakarta Malay *di-*) has been lost in SLM, as it was in Moluccan varieties such as Ambonese Malay.

41. The nasal coda is often deleted, producing the form *na*. The onset of the full form is denasalized following monosyllabic pronouns ending in a vowel: *sɛ -daŋ* (1s), *dɛ-daŋ* (3s), *go-daŋ* (1s), but *dia-naŋ* (3s).

(29) *guruvərəya-ʈə hondə pæ:na-k tiyɛ-nəwa.* SINHALA
teacher-DAT good pen-INDEF EXIST-PRES
'The teacher has a good pen'

(30) *guru-naŋ ada baik pena.* SLM
teacher-DAT EXIST good pen
'The teacher has a good pen' (Saldin 2001:66)

The object marking function of *naŋ* can be regarded as isomorphic to the Tamil dative inflection exemplified in (26); however *naŋ* can also mark clausal complements which require an accusative case affix in Tamil. In Tamil, the case-marking of nominalized clauses is determined by the syntactic status of the nominalized clause in relationship to a predicate. Tamil nominalization affixes are functionally and formally distinct from case affixes.[42] Although *naŋ* is a reflex of dative case on object DPs, in nominalized clauses right-peripheral *naŋ* does not reflect the thematic status of the clause (nor of DP arguments within the clause) in any way.[43]

5.4 Pre-nominal relative clauses

The SLM relative construction resembles Tamil and Sinhala relatives in that the nominal head invariably follows the relative clause. This is utterly unlike what we find in Malay dialects spoken outside of Sri Lanka. Yet according to Arden (1969:203), "there are no relative pronouns in Tamil. Their place is supplied by participles with a relative meaning". The corresponding SLM construction, in spite of the equivalent surface distribution of lexical constituents, contains the 'relative pronoun' *nya*, according to Saldin (2001:64–65), a judgment with which other native speakers I have consulted concur. Saldin distinguishes between an adjectival construction in SLM, lacking a relativizer, and the actual relative construction in that language, which always contains a relativizer. The Tamil and Sinhala relative constructions are the linear equivalent of the SLM pre-nominal adjectival modifier. The adjectival modifier is simply the relative

42. The following example from Lehmann (1989:302) illustrates the Tamil facts:

(i) [*kumaar maturai-kku poo-n-at-**ai***] *raajaa so-nn-aan.*
Kumar Madarai-DAT go-PAST-NOMINAL-ACC Raja say-PAST-3SM
'Raja told (the fact) that Kumar went to Madurai'

The embedded clause is the nominal complement of the matrix verb, and consequently receives overt accusative case, which is spelled out on the nominalized embedded verb. Lehmann adds that "nominalized clauses occurring as object arguments to the above mentioned type of verbs are always interpreted as factive complements, that is as true propositions... a false or untrue proposition cannot be embedded as a nominalized clause."

43. Saldin (1996) speculates that the *naŋ* in embedded clauses may be etymologically derived from the Malay preposition *dengan* (=with). It is more plausible that *naŋ* is derived from the homophonous Javanese preposition, as stated above.

construction with no relativizer. The Tamil and Sinhala relatives are explicitly marked as adjectival, which is not an option either for the adjectival or the relative construction in SLM. The sentence in (31) from Saldin (1996:65) is an example of the SLM relative construction.

(31) [[***jalan ka nya lari***] ***ayaŋ***] *baisikal atu ka tərbuntur su mati-mati.*
road P REL run chicken bicycle DET P struck PAST die
'The chicken which ran along the road was struck by a bicycle and died'

A still more radical contrast with Tamil (than the fact that prenominal relative constructions contain an overt relativizer) is the fact that the relativizer invariably precedes the verb within the relative clause. In Tamil clauses, any complementizer-like element appears at the right edge of the clause. In declarative complement clauses in SLM, similarly, a tensed verb does occur to the left of a quotative particle, as in (32). Tense and negation markers cannot co-occur, so *pi* would be unmarked for tense if it were negated.

(32) *Ali tər-tau* [[*kakɛ məsigit naŋ su pi*] ***kətə***]. SLM
Ali NEG-know grandfather mosque P PAST go QUOT
'Ali doesn't know that grandfather went to the mosque'

In the Tamil translation of this sentence, the embedded clause is center-embedded. Dravidian languages, including Tamil (Bayer 1998), never have both initial and final complementizers. The SLM complementizer *kalo* 'if' occurs variably in clause-initial and clause-final position, which suggests competition between Lankan and (older) Malay grammatical options.

6. An external explanation for the domain contrast?

It is quite likely that an early variety of SLM was strongly influenced by the grammars of L2 speakers who acquired Malay in the context of mixed marriages. We know from the work of the SLM historian Hussainmiya (1990:48) that Malay/'Moor' intermarriage was rather frequent in the early period, and it is still rather frequent due in particular to shared religious beliefs and institutions. Without extensive L2 acquisition of Malay by speakers of Tamil, it is difficult to imagine how such extensive typological change could have taken place in the space of three hundred years. The bilingual family is the sociolinguistic domain that would have been most conducive to grammar convergence. This would help to explain the impressive extent of convergence of SLM on the grammar of Muslim Tamil (mediated by the grammar of L2 Malay); however it would not explain the cross-categorial asymmetry in the extent of that convergence. The apparent surface similarities between the syntax of DP in SLM and DP in Tamil rarely reflect the grammar of the lexical source language of SLM (Malay). On the other hand, they can also not be exclusively attributed to 'substrate influence' (as opposed to convergence of L1 and L2 varieties), unless the grammar of SLM had

been so strongly influenced by Muslim Tamil-speaking mothers who transmitted an L2-influenced Malay variety to their children that the original L1 contact variety was completely submerged. A historical hypothesis associating SLM grammar exclusively with the ostensible L2-speaker status of a majority of SLM-speaking mothers[44] would also not account for the conservative pre-verbal distribution of temporal elements and negation. The verb data falsify any claim that the grammar of SLM is either the product of simple regrammaticization of Malay or of simple relexification of Tamil. The morphosyntax of SLM is better characterized as a compromise between contact Malay and the L2 variety of earlier Muslim Tamil speakers, with adstratal influence from Sinhala due to later Malay-Sinhala bilingualism in the SLM-speaking population. Even if the verbal domain had converged on the grammar of Tamil to the extent that the nominal domain has, there is no one-to-one cross-linguistic mapping of the grammatical properties of lexical items. From a recent generative perspective according to which syntactic structure determines morpheme order, the feature content of lexical items should determine morpheme order within phonological words, including SLM verbs (and nouns). I take a contrast in distribution between a grammatical source language (Tamil) and a contact language (SLM) as evidence that the features of lexical items and functional elements in the two languages differ. Even from a purely descriptive vantage point, SLM verbs are still morphologically unlike Tamil verbs in their linear organization (and radically unlike those of Sinhala). In spite of the relatively Lankanized nominal domain in SLM, considerable surface differences remain between

44. Detailed historical demographic research must precede a categorical claim that SLM is exclusively the product of nativization of the Malay L2 grammars of native Tamil speakers. Although there has been considerable intermarriage between Malays and Muslim Tamil speakers in Sri Lanka, and this appears to have played a major role, the precise historical demography is unknown. There have been references to single and married females accompanying male migrants in the colonial period, and the details of these migrations are in need of further investigation.

Jayah (1970), for example, writes that "Bahasa Melayu (the Malay language) has been preserved in this country for over 250 years due to the fact that the original exiles from Indonesia were accompanied by their womenfolk and it was *not necessary* for them [my emphasis] to find wives among Sinhalese and Tamil women, unlike the Arab ancestors of the Ceylon Moors."

A view that an unbalanced ratio of male to female immigrants *presupposes* that the language of the household in a majority of linguistically-mixed marriages was *necessarily* Malay is challenged by anthropological research pointing to quite a different pattern, at least for parts of Sri Lanka. At least in the twentieth century, it was more common in mixed marriages for the mother's L1 to be transmitted to children, even when that language was Tamil rather than Malay (M. Jaffar, p.c.). A husband was absorbed into his wife's extended family. Anthropological research (McGilvray 1989) on matrilocal domestic social organization in Tamil-speaking communities (both Hindu and Muslim) challenges the view that Tamil-speaking women necessarily transmitted L2 Malay to children. A Malay husband, linguistically isolated in his wife's Tamil-speaking extended family, could not easily establish a Malay-speaking environment in the domestic sphere. Domestic language selection in Colombo may have been more variable however and more data on earlier patterns will be needed.

the major Lankan languages and SLM, given an un-Lankan relativizer (*nya-*), case morphology which strongly underdetermines the case system of Tamil, and variable object movement which does not encode contrastive focus.

7. The development of SLM

The language's development has involved the evolution from an ordinary contact Malay variety to what may be characterized as a compromise between the grammars of such varieties and those of the dominant Lankan languages. SLM now features clause types which are more Lankan than Malay. These are associated with the nominal domain, and follow from the generalization that nouns are modified on the left (right-headed relatives), and that DPs raise to the left within PP (non-finite nominalizations). Although SLM differs from other Malay varieties in marking verbs as non-finite, the apparently clausal constructions in which this takes place contain a nominalized VP. The presence of OV clauses in all lects is a function of object movement. Key changes to the nominal syntax of SLM from what we find in other Malay varieties can be schematized as follows:

1. obligatory object movement within PP (DP-P order) renders prepositions – including *naŋ* – postpositional
2. *na*(*ŋ*) (=P) functions as a reflex of dative case on nouns, including the complements of postpositions, while *nya*(*ŋ*) or *yaŋ* (depending on the dialect) marks accusative case
3. case-marked DP objects raise to the left of the matrix verb
4. *naŋ* takes a nominalized clausal complement containing non-finite VP, assigning invariant case to the nominalization
5. case-marked nominalizations raise to the left of the matrix verb

There is an implicational relationship between 1 and 2 and between 1 and 4. The relationship between 1 and 2 is paralleled in the relationship between 1 and 4. Case-marking and other apparently inflectional marking of nouns in SLM are uniformly right-peripheral to the noun itself, as in Tamil and Sinhala. Since this is so, and since the greatest lexical resource for case-marking has been the set of contact Malay adpositions (with the exception of *nya*(*ŋ*)/*yaŋ*), the reorganization of PP had to take place prior to the functional expansion of adpositions as case markers.[45]

In language contact scenarios, we expect that change does not take place in all parts of a grammar simultaneously. A single change may precipitate a series of con-

45. I prefer the term *functional expansion* to *reanalysis* here, since case markers such as *naŋ* retain their adpositional status, with functional status dependent on syntactic context. Within certain PPs, *naŋ* is not a P head, but a secondary marker of dative (oblique) case on the noun phrase. In other contexts, *naŋ* is difficult to distinguish from an ordinary postposition. Certainly there is no phonological motivation for treating it as a case affix.

sequences. For SLM, several critical developments have included object movement within PP yielding invariant post-nominal adposition orders, the case-marking of object DPs, and the nominalization of VP. Subsequently, both case-inflected DP objects and case-inflected nominalizations acquired the ability (and later the propensity) to raise to the left of a finite verb. This is the grammar in which *finite* clauses are *also* OV, as in Tamil and Sinhala. Given that the relationship between temporal affixes and the finite verb appears to be the same in all varieties, it does not seem appropriate to posit a verb movement contrast. Rather, the contrast between matrix SVO and matrix SOV is *also* a function of changes in the nominal domain, with the verb system stable across SLM grammars. This analysis provides a theoretical approach to a part of the diachrony on which there has been no prior grammatical analysis, and about which unfortunately far too little is known, owing to a lack of vernacular texts from before the twentieth century.[46]

8. Conclusion

The fact that SLM has evolved a set of pre-verbal, albeit phonologically-dependent, functional markers of verbs, and the fact that the SLM verb does not permit the stacking of these morphemes at the verb's left edge to which most are restricted in their distribution, make it difficult to see how the verbal domain and its functional properties can be treated as straightforwardly Tamil-derived. The adjacency requirement does reflect Tamil influence, but the convergence process is nevertheless incomplete. In the case of tense-marking, the development of that morphological process through functional reanalysis is indeed likely to have been a product of Malay-Tamil bilingualism; however, again it follows from the pre-verbal distribution of tense markers that convergence is incomplete. Sinhala tense morphology, relying heavily as it does on *ablaut* and other phonological alternations, bears no resemblance at all to what we find in SLM. Other morphosyntactic properties of the SLM verb can be attributed to the Malay antecedent lects for which SLM is a Lankan-influenced *koinē*. SLM aspect markers, although post-verbal, do not sufficiently reflect the syntax of aspect in Tamil to constitute a straightforward mapping from that language to SLM. Furthermore, the overt encoding of aspectual contrasts – albeit with pre-verbal particles – is robustly attested in pre-migratory Malay dialects, so we should not view the SLM aspectual system as simply mapping from the Tamil system, nor as necessarily the product of creolization. In conclusion, the current extent of morphological and syntactic convergence with Tamil and Sinhala, although striking, is incomplete, and is asymmetrical across syntactic domains. The non-finite clauses in SLM point to Sinhala influence; however, as nominalizations they support the generalization that it is primarily nom-

46. Nineteenth-century literature was not written in spoken SLM, but in a classicizing literary variety.

inal constructions in SLM that align with their counterparts in Tamil and/or Sinhala. The development of internal and external accounts that conclusively explain the asymmetry in the extent of convergence in the nominal and verbal domains awaits future research.

References

Adelaar, K. A. (1991). Some notes on the origin of Sri Lanka Malay. In H. Steinhauer (Ed.), *Papers in Austronesian Linguistics* No. 1 [Pacific Linguistics, Series A-81] (pp. 23–37). Canberra: The Australian National University, Research School of Pacific Studies.

Ansaldo, U. (2005). Typological Admixture in Sri Lanka Malay. The Case of Kirinda Java. Manuscript. Universiteit van Amsterdam.

Arden, A. H. (1969). *A Progressive Grammar of the Tamil Language*. Chennai: The Christian Literature Society.

Bakker, P. (2000). Convergence intertwining: An alternative way towards the genesis of mixed languages. In D. Gilbers, J. Nerbonne, & J. Schaeken (Eds.), *Languages in Contact* [Studies in Slavic and General Linguistics 28] (pp. 29–35). Amsterdam: Rodopi.

Bayer, J. (1998). Final Complementizers in Hybrid Languages. Manuscript. Friedrich-Schiller-Universität-Jena.

Bichsel-Stettler, A. (1989). Aspects of the Sri Lanka Malay Community and its Language. M.A. thesis, Universität Bern.

Hussainmiya, B. A. (1986). 'Melayu Bahasa': Some preliminary observations on the Malay Creole of Sri Lanka. *Sari, 4*(1), 19–30.

Hussainmiya, B. A. (1990). *Orang Rejimen: The Malays of the Ceylon Rifle Regiment*. Bangi: Penerbit Universiti Kebangsaan Malaysia.

Jayah, M. (1970). The plight of the Sri Lankan Malays today. In *MICH Silver Jubilee Souvenir, 1944–1969*. Colombo: Moors' Islamic Cultural Home.

Lehmann, T. (1989). *A Grammar of Modern Tamil*. Pondicherry: Pondicherry Institute of Linguistics and Culture.

McGilvray, D. (1989). Households in Akkaraipattu: Dowry and domestic organisation among the matrilineal Tamils and Moors of Sri Lanka. In J. N. Gray & D. J. Mearns (Eds.), *Society from the Inside Out: Anthropological Perspectives on the South Asian Household* (pp. 192–245). New Delhi: Sage Publications.

Robuchon, G. (2003). Malayo Language in Sri Lanka. Paper presented at the 7th International Symposium on Malay/Indonesian Linguistics in Nijmegen.

Saldin, B. D. K. (2001). *The Sri Lankan Malays and their Language* (Revised edition). Kurunegala: T.S. Jamalon.

Slomanson, P. (2001). A Malay substrate 'passive' model for contact Dutch. Paper presented at the Society for Pidgin and Creole Linguistics meeting at the University of Coimbra, Portugal.

Slomanson, P. (2005). The Verbal Morphosyntax of Non-Canonical Contact Languages – Malay-derived Constraints and the Inflectional Domain in Afrikaans and Sri Lankan Malay. Ph.D dissertation. City University of New York Graduate Center.

Smith, I., Paauw, S., & Hussainmiya, B. A. (2004). Sri Lanka Malay: The state of the art. In R. Singh (Ed.), *Yearbook of South Asian Languages* (pp. 197–215). Berlin: Mouton de Gruyter.

Steinhauer, H. (Ed.). (1991). *Papers in Austronesian Linguistics* No. 1 [Pacific Linguistics, Series A-81. Canberra]. Canberra: The Australian National University, Research School of Pacific Studies.

van Minde, D. (1997). *Malayu Ambong: Phonology, Morphology, Syntax*. Leiden: CNWS Publications.

Sri Lanka Malay

Creole or convert?[1]

Ian Smith and Scott Paauw
York University, Canada

Two sociolinguistic characteristics of Sri Lanka Malay (SLM) are atypical for a creole: a non-European, non-colonial lexifier (Vehicular Malay), and a single substrate (Sri Lanka Muslim Tamil) with which it remained in contact. The Tamil-like characteristics of present-day SLM differ from those often claimed to typify creoles. In particular, the SLM tense-mood-aspect system expresses Tamil categories rather than the traditional creole categories of anterior, nonpunctual and irrealis. Bakker (2000) claims 19th century SLM more closely resembled a 'typical' creole, and the Tamilization of SLM represents a recent process of 'conversion'. However, present-day SLM descends from (unrecorded) 19th century colloquial SLM, not from the written diglossic high. The Tamilized structure of SLM results more plausibly from its atypical developmental context than from recent influence.

---

1. We owe a great debt of gratitude to B.A. Hussainmiya for providing much of the data on which this paper is based and for answering our many questions. We acknowledge the support of the Social Sciences and Humanities Research Council and the Faculty of Graduate Studies of York University for grants which supported some of our research. We have also benefited from comments by Sander Adelaar, Peter Bakker, James Gair, Peter Slomanson, anonymous reviewers and other colleagues. Any errors are of course our own. Abbreviations used in glossing are as follows: ACC: accusative; ADVZ: adverbializer; AGR: agreement; COLLOC: collocative; COND: conditional; CONJ: conjunction; DEB: debitive (obligatoriness); DAT: dative; EB: elder brother; EMPH: emphatic; FB: father's brother; FOC: focus; FUT: future; GEN: genitive; HABIL: habilitative (ability); HORT: hortative; INF: infinitive; LINK: linker; LOC: locative; NEG: negative; NOM: nominative; NOMZ: nominalizer; PFV: perfective; PL: plural; PPL: participle; PRF: perfect; PRES: present; PROG: progressive; Q: question marker; QUOT: quotative; REP: reportative; VOL: volitive; YB: younger brother; 1SG: 1st person singular pronoun (etc.). Transcriptional conventions differ slightly for each language: for Standard and Vehicular Malay, we use Malay orthography with the addition of ĕ to indicate schwa; for Sri Lanka Malay, the same conventions are used with the addition of IPA symbols for retroflexes and doubled vowels to indicate phonemic length; for Tamil we use IPA modified by South Asianist conventional usage of <y> for [j], <c> for [tʃ] and <j> for [dʒ];

1. Introduction

The majority of known creoles arose from European colonial activities, and have European lexifiers; most also have several substrate languages. Common structural characteristics of creoles may, in theory, result from the nature of the creolization process, the structural similarity of European lexifiers, similarities in substrate languages, or the social context of creolization (e.g., differences between fort and plantation situations). To gain a wider perspective on the processes and results of creole genesis it is important to investigate a language such as Sri Lanka Malay (SLM), whose lexifier (Vehicular Malay) is neither European nor the language of a colonial power and which has a single substrate (Sri Lanka Muslim Tamil [SLMT]).

Several authors have noted the strong influence of Tamil on SLM (Hussainmiya 1986; Adelaar 1991; Smith 2003). Many of its characteristics run counter to those often claimed to be typical of creoles. This paper focuses on tense, mood and aspect (TMA), showing in particular that the TMA system of SLM expresses Tamil categories of tense (PRESENT, PAST, FUTURE) aspect (PERFECT, PERFECTIVE, PROGRESSIVE, UNMARKED) and mood (VOLITIVE, DEBITIVE, HABILITATIVE, etc.) rather than 'traditional' creole categories (ANTERIOR tense, DURATIVE aspect and IRREALIS mood, each contrasting with a semantically and formally unmarked category). We also argue that these Tamilized TMA categories in SLM developed as a result of the creolization process rather than by recent contact-induced 'conversion' as suggested by Bakker (2000a, 2000b).

The paper is organized as follows: first some background socio-historical information on SLM; second, a brief discussion of TMA in Malay; third a comparison of TMA in SLM and Sri Lanka Muslim Tamil; fourth, discussion and conclusions, including arguments for an early, rather than a late, development of the TMA characteristics of SLM.

1.1 Origins of Sri Lanka Malay

SLM is spoken by descendants of people brought to Sri Lanka from the Indonesian archipelago and the Malay peninsula by Dutch and British colonial administrations. Most of these immigrants were soldiers, with smaller numbers of political exiles, slaves, and convicts. From their arrival, they associated closely with the established Tamil-speaking Sri Lanka 'Moor' community, with whom they shared the Muslim religion (Hussainmiya 1987: 45). As the soldiers were single males for the most part, a high degree of intermarriage with the Moor community was inevitable. Through this interaction between the two communities, a new language arose with Tamilized structure and with a Malay lexicon.

---

allophonic voicing in the stops is spelled out to facilitate pronunciation for those unfamiliar with the language.

The first waves of immigrants brought by the Dutch, beginning in 1656, came from various locations in the Indonesian archipelago, and were native speakers of a number of West Austronesian languages.[2] As most were military recruits, they had received exposure to varieties of Vehicular Malay used as lingua francas by the Dutch in the East Indies. All recruits spent time in Batavia (Jakarta) prior to departure, and there is some evidence of possible influence of Jakarta Malay[3] on Sri Lanka Malay. A form of Vehicular Malay became the lingua franca of the East Indies community in Sri Lanka, and eventually led to the community identifying itself as 'Malay', an identity reinforced by the British administration, beginning in 1796, which recruited Malay soldiers largely from the Malay peninsula.

Today, although the Malay community in Sri Lanka numbers approximately 48,000 out of a population of 16.8 million (Sri Lanka 2001 Census), SLM is spoken by a dwindling number of people. Traditionally Malay men were bilingual in Tamil (through their religion) and, after the British took control of the island, English (through their connection to the military); knowledge of Sinhala was likely less widespread. Malay was maintained because women were often monolingual and because the predominantly Malay Ceylon Rifle Regiment (CRR) provided a focus for the community as well as instruction in Standard Malay (Hussainmiya 1990). Bichsel-Stettler (1989) documents how SLM has gradually lost its domains of use in Sri Lanka: the CRR was disbanded in 1873, regimental schools were abandoned and literacy declined; from the 1930s, female participation in the educational system increased, producing more bilingualism among women; also from the 1930s, Tamil began to replace Malay in mosques and other religious contexts; after independence in 1948, when Sinhala gained importance for government employment, and particularly after 1956 when English-medium education became restricted to private schools, many Malay families opted to send their children to Sinhala medium public schools, but those who were able often began speaking English at home in order to pass it on to their children. Thus among the middle and upper classes of Colombo studied by Bichsel-Stettler, Malay is being lost in favour of Sinhala/English bilingualism. In outlying areas, such as southeastern Sri Lanka, the maintenance of SLM/Tamil bilingualism is, however, considerably stronger (U. Ansaldo, p.c. May 2004). P. Slomanson (p.c., Oct. 2005) points out that SLM/Tamil bilingualism is maintained where children attend (Tamil medium) Muslim schools. The largest SLM communities are found in Colombo, Kandy, Badulla and Hambantota.

---

2. Immigrants to Sri Lanka brought by the Dutch were from Ambon, Banda, Bali, Java, Madura, and Buginese and Malay areas (Adelaar & Prentice 1996: 685).

3. The fact that the Betawi community (descendants of the original Malay speaking residents of Jakarta) identifies itself as 'Malay' may also have influenced the self-identity of the Sri Lanka Malay community.

1.2 Vehicular Malay

We use the term *Vehicular Malay* to avoid the confusion caused by the more familiar terms *Bazaar Malay* (which can refer to a pidgin spoken in the Malay peninsula or to the trade language that gave rise to contact varieties of Malay throughout the archipelago) and *Low Malay* (which can refer to the colloquial form of 'native' varieties of Malay or to possibly creolized contact varieties of Malay). *Vehicular Malay* (henceforth VM) refers to varieties of Malay used as lingua francas in contact situations, without prejudging their possible status as pidgins or creoles. VM, in the Sri Lankan case, refers to the varieties of Malay brought by immigrants which lexified SLM. Varieties of VM[4] share a set of common features to a greater or lesser degree (Adelaar 1991; Adelaar & Prentice 1996: 675; Paauw 2003, 2004, 2005): e.g., merging of word-final nasals, the possessive marker *pe*, a second person pronoun derived from Hokkien (*lu*), loss of schwa (in some dialects), and loss of word-final *-h*.

1.3 Sri Lanka Malay

Unlike VM, which features simplified Malay morphosyntax, SLM has developed, through the influence of Tamil, a complex morphosyntax which differs greatly from any other variety of Malay. Some features of SLM not seen in other varieties of Malay are postpositions, SOV word order, and case markers, all of which are common in Tamil (and Sinhala, the majority language of Sri Lanka). SLM, in its morphosyntax, more closely resembles Tamil and Sinhala than it does Malay, a characteristic it shares with another creole language of Sri Lanka, Sri Lanka Portuguese (Smith 1978, 1979a, 1979b). Because of the similarities between Sinhala and Tamil, distinguishing the influence of one from that of the other is not always possible. Where distinctions can be made, Tamil appears to have been the dominant influence (Smith, Paauw & Hussainmiya 2004; Smith 2003) as would be predicted from the close contacts between Malays and Moors. Although SLM has retained a predominantly Malay lexicon, the morphosyntactic developments have made it incomprehensible to speakers of other varieties of Malay (which are, themselves, largely mutually comprehensible). The phonemes of VM all appear in SLM. Retroflex and geminate consonants also occur (through substrate influence), though their status is phonologically marginal; one clearly phonemic substrate-influenced characteristic is distinctive vowel length. A structural sketch of SLM can be found in Smith, Paauw and Hussainmiya 2004.

Spoken SLM exhibits considerable variability, a characteristic to be expected, given the diverse origins of the Malay speakers brought to Sri Lanka, the wide geographical dispersion of the community, and the fact that it has never been written. For much

4. Other varieties that could be considered to be descended from Vehicular Malay include those of Manado, Ambon, North Maluku, Larantuka, and Kupang in Indonesia, as well as Baba Malay in Malaysia and Singapore, Melayoe Sini in the Netherlands, and historical varieties, such as the forms termed *Tangsi Malay* and *Java Malay* by Adelaar and Prentice (1996: 678–679).

of the 19th century, Standard Malay functioned as a diglossic high in the community, a fact which will become relevant in the discussion section below. Even today, some members of the community have become familiar with Std. Malay, and some published sources of SLM written by community members also exhibit Std. Malay influence (e.g. Saldin 1996). The primary data used for this paper are taken from the recording of a natural conversation made in Kirinda, a village in southeastern Sri Lanka, in the late 1970's by community member, Dr. B. A. Hussainmiya, and subsequently transcribed and translated into Sri Lanka Muslim Tamil with his assistance. The transcript consists of approximately 500 mostly sentence-length items. These have been supplemented by elicitations and data taken from other sources.

The focus of this paper is on TMA in SLM. We will consider various possible sources of TMA marking in the language: the superstrate, VM; universals of the creolization process; and the primary substrate language Tamil.

2. TMA in Vehicular Malay

In VM, as in Std. Malay (also known as *Literary Malay* or *High Malay*), tense is unmarked, with time reference being inferred from context. Mood and aspect are indicated by discrete markers, referred to variously as adverbials, pre-modifiers, aspect markers and temporal markers. These markers always precede the predicate, are not bound, and can occur with other markers of this set, within certain parameters.[5] Examples are seen in (1)–(3).

(1) *Ali sudah bangun*
Ali PFV wake
'Ali has woken up' (Std. Malay, Sneddon 1996: 197)

(2) *Se masi balong guru lei*
2SG STILL NOT.YET teacher ALSO
'You're still not a teacher yet' (Ambon Malay, van Minde 1997: 211)

(3) *Kami harus sudah bisa berdiri sendiri tahun ini*
1PL DEB PFV HABIL stand alone year this
'We should have already been able to stand alone this year' (Std. Malay, Alwi et al. 1998: 160)

5. One reason that varieties of Malay other than SLM do not show significant variation in their syntax or ways of indicating TMA is that the substrate languages involved (usually Austronesian languages, or, in the case of Baba Malay, Hokkien) are very similar syntactically to Malay, and any substrate influence would thus be covert.

3. TMA in SLM, Tamil and VM

We now outline the main characteristics of TMA marking in SLM in comparison with (Sri Lanka Muslim) Tamil and VM. As we shall see, SLM and Tamil are largely congruent in the categories that they mark, but not necessarily congruent in terms of their morphosyntactic expression of these categories. This observation also holds for Sri Lanka Portuguese (Smith 1979a) and is not unexpected. An overview of the system is seen in Table 1; the details are presented in the sections below. Conventions used in the table are as follows: affixes/clitics are indicated by a hyphen (–) and auxiliaries by a plus sign (+),[6] these symbols being placed after prefixes and auxiliaries that precede the verb and before suffixes and auxiliaries that follow the verb. A slash (/) separates allomorphic variants; a semi-colon (;) separates alternate expressions. Forms written with no intervening slash or semi-colon occur together.

3.1 Tense

SLM has three primary tense categories, PRESENT, PAST and FUTURE, marked by verbal proclitics. These same categories are found in Tamil, where they are marked by suffixes. In SLM tense markers are normally procliticized to the verb (and thus look like prefixes) but in very rare instances they seem to procliticize to a preverbal adverb or other element, as in (8) below. Their preverbal position in SLM reflects the fact that they mostly derive from VM preverbal TMA markers.

3.1.1 *Present: ara-*

Present tense can be used to mark events and states situated around the present time, broadly as in (4) and (5), or more narrowly as in (6) and (7), as well as future actions viewed as prearranged or certain, as in (8).

(4) SLM maana-ka kuuli pĕrajang ara- kĕrja
SLMT enge kuuli ʋeele sey -r -aanga
where LOC daily.wage work PRES do PRES AGR
A: 'Where do they do daily wage work?'
SLM uuḍik-paḍa ara- pii
SLMT naaʈʈup.pakkam-gaḷ poo -r -aanga
country.area-PL PRES go PRES AGR
SLM siitu siini ara- kinja
SLMT ange inge sey -r -aanga
there here PRES do PRES AGR
B: 'They go to country areas; they work here and there.'

6. In typologically postposing languages, the distinction between case suffix and postposition can be fuzzy.

Table 1. TMA markers in SLM in comparison with SLMT & VM

	SL Malay	SL Muslim Tamil	Vehicular Malay	Std. Malay cognates
Tense				
Present	ara-	-(kk/g)-r-/-(kk/g)-	not marked	*ada* 'be' (existential)
Past	ay(ng)-/nya(ng)-; su(ɖa)-/sa-/sĕ-	-n-/-tt-/-d-/-nd-	not marked	??; sudah (PERFECTIVE)
Future	ati-; ang-	-ʋ-/-pp-/-kk/g-	*nanti*+ (optional)	*nanti* 'later, wait'
Aspect				
Perfect	+aaɖa	+iru	not marked	*ada* 'be' (existential)
Perfective	+abis; +taaro (etc.?)	+(ʋi)ɖu; +pooɖu etc.	*suda*+	*habis* 'finished'; *taruh* 'put'
Progressive	+ambel +aaɖa	+koɳɖu +iru	*lagi*+ ; *ada*+	*ambil* 'take' *ada* 'be' (existential)
Mood				
Volitive	N-DAT … (ma-)V(-na) +maaʋ	N-DAT … V-INF +ooɳum	*mau*+	*mau* 'want'
Debitive	N-NOM … (masi-/ma-/mang-)V +maaʋ	N-NOM V-INF +ooɳum	*mes*(*t*)*i*+	*mĕsti* (DEBITIVE)
Habilitative	(mang-)V-na +boole	V-INF +eelum	*bole*+ ; *bisa*+	*boleh* 'permission'
Reportative	+kanyang	+aam	*kayanya* ; *katanya*	katanya (?)

(5) SLM sepe muuɖu joo ara- jaaɖi
SLMT eeɳɖa aacca daan aaʋ -r -aar
1SG.GEN FB FOC PRES become PRES AGR
'He happens to be [lit. becomes] my father's brother'

(6) SLM kitang diya-na atu le tira kinja kaʈa
SLMT naanga aʋar-ukku oɳɳ um seyya illa eɳɖu
1PL 3SG-DAT one EMPH NEG do NEG QUOT
SLM ara- iingat
SLMT nene -kk.r -aar
PRES think PRES AGR
'He thinks we won't do anything to him' [Context: Subject is alleged to have stolen speaker's boat.]

(7) SLM ii.pii.ef le ara- pootong deram.paɖa-pe
SLMT ii.pii.ef um ʋeʈʈ -r -aanga aʋan.ga-ɖa
EPF EMPH PRES cut PRES AGR 3.PL-GEN
'They are cutting off the E[mployers'] P[rovident] F[und] too – of those guys'

(8) SLM see ara- konyam blaajar
SLMT naan konyam ʋaasi -kk.r -an
1SG PRES some read PRES AGR
'I'll just read some…' [Then speaker begins to read.]

The marker *ara-* derives from VM. Adelaar (1991:31) points out that it serves as a progressive marker in Moluccan, Bazaar and Baba Malay (1991:26) and that it derives ultimately from the existential verb *ada* (cf. also Saldin 1996:62). In SLM *ara-* is not specifically progressive (though glossed as such in Bichsel-Stettler's (1989) texts), but mirrors the various functions of the Tamil (and Sinhala) present. A contrastively progressive formation is discussed in 3.2.3. The form *aaɖa* continues in SLM as an independent quasi-verb meaning 'be', 'exist' and as the perfect auxiliary. (See 3.2.1)

3.1.2 *Past: ay(ng)-/nya(ng)-; su(ɖa)-/sa-/sĕ-*

There are two main past markers in our data which likely derive from competing markers in the VM of speakers of different origins. One marker, *su(ɖa)-/sa-* deriving from the VM PERFECTIVE marker *suda*, is well known in the literature (Saldin 1996:63; Robuchon 2003:12) and is found in other varieties of Malay. (*Su* is also found in East Indonesian varieties of VM.) It is glossed as PERFECTIVE in Bichsel-Stettler's texts (1989), but the contexts do not support this label. (See 3.2.2 for the SLM PERFECTIVE.). The full form *suɖa* also occurs in SLM as a discourse particle equivalent to SLMT *ini* 'now', 'so', 'here', seen in (9). A second past marker, of obscure etymology, appears as *ayng- or nyang-* or various realizations of these: the final nasal assimilates to a following consonant and is often realized only as nasalization (thus [ɐ̃j̃] or [ɲɐ̃]); often it disappears completely ([ɐj] / [ɲɐ]); we have also recorded *ang-* and a syllabic

nasal. This marker is also found in Bichsel-Stettler's texts, but is not noted by Saldin (1996). It does appear, however, in one of his examples, repeated here as (10). He also gives *ai-* as a southern variant of *asa-*, which he misanalyzes as a simplification of the PERFECTIVE marker, habis (1996: 63). No distinction apparently exists between *ayng-/nyang-* and *su*(ɖ*a*)-; in (11)–(15) naturally occurring examples of various past-marked forms of *pii* 'go' are shown. Many 'conjunctive'[7] past participles in our data have double marking, always in the order *ayng-su-/nyang-su-*, which can contract to as(i)-, as in (23) and (33), and this is doubtless Saldin's *asa-*. Indeed, his example seen in (10) involves a conjunctive past participle. In Bichsel-Stettler's texts the prefix *as-* (glossed as IMPERFECTIVE) also marks conjunctive past participles.

(9) SLM suɖa kitang ara- biilang tuʋan aapa -na
SLMT ini naanga sol -r -oom tura enna eɳɖ-aa
LINK 1PL PRES say PRES AGR sir what -DAT QUOT-COND
'So what we are saying, sir, is. . . .'

(10) SLM se kantor-nang asa- pi rumah-nang nya- datang
1SG office-DAT PAST.PPL go house-DAT PAST go
'I went to the office and came home' (Saldin 1996: 63, interlinear gloss added)[8]

(11) SLM raazik fariid, derang tembak -na su- pii
SLMT raazik fariid, aʋanga sud -a poo -n -aanga
R F 3PL shoot INF PAST go PAST AGR
'Razik Fareed, they went to shoot [him]'

(12) SLM atu skaali m- pii
SLMT oru taɖaʋe poo -n -aanga
one time PAST go PAST AGR
'At a certain time [they] went'

(13) SLM sĕpaaru oorang lompat nya- pii
SLMT araʋaasi peer paa -nj -i poo -n -aanga
half people jump PAST -PPL PAST go PAST
'Half the people escaped' [lit: . . .having jumped, went]

(14) SLM pəmisyan tĕra.ana aym- pii, derang
SLMT pəmisyan ill.aama poo -n -aanga, aʋanga
permission without PAST go PAST AGR 3PL
'[They] went without permission – them'

7. The South Asian areal feature known most commonly as the 'conjunctive participle' is an indeclinable verb form used to link two clauses which normally share a subject and often other elements. See Masica (1976: 108–140).

8. We retain Saldin's original orthography, which does not distinguish long vowels or retroflex consonants.

(15) SLM mari pii kaṭa aym -pii
SLMT ʋaanga poo -ʋ -oom eṇḍu poo -n -oom
HORT go FUT AGR QUOT PAST go PAST AGR
'Saying 'let's go' [we] went'

A third marker, the suffix *-le*, is reported by Robuchon to be "the most largely used pattern for the past" by the Colombo-area youth who served as his main informant (2003:2). In our data, *le* (< Malay *lagi* 'also, yet, again') occurs only as a conjunction/emphasis marker parallel to Tamil *um*, seen in (6), (7), (16), (25), (37).

3.1.3 *Future: ati-/ang-*

The future occurs infrequently in our corpus. Two markers appear, the first, *ati-*, derives from the VM adverbial *nanti* and is cognate with the Std. Malay (ad)verb *nanti* 'later; wait' (Saldin 1996:63); the second future marker, *ang-*, is likely cognate with the Std. Malay future marker *akan*. Again the variation between these two forms reflects the disparate origins of the speakers of VM. Examples are seen in (16)–(17). Saldin (1996:63) also reports the form *anti,* and Robuchon (2003:14) has *ti* in his corpus; both are obviously also reduced forms of *nanti*. It should be noted that while Sinhala has present and past categories that parallel those of Tamil and SLM, it marks future time modally through volitive or involitive optatives (Gair 1976). VM can also mark future time modally through the use of the volitive marker *mau.*

(16) SLM maana ʋaktu le aapa ati- jaaḍi kaṭa ta-
SLMT enda neeram um enna naḍa -kk -um eṇḍu
which time EMPH what FUT happen FUT- AGR QUOT NEG
SLM taaʋ
SLMT teriy -aa
know FUT.NEG
'At any time, we don't know what will happen'

(17) SLM kaarang deram.paḍa-pe nigrii oorang-paḍa kaakaa aaḍe
SLMT ippo aʋanga-ḍa uur aaḷ-kaḷ kaakaa tambi
now 3PL-GEN village person-PL EB YB
SLM kaṭa
SLMT eṇḍu
QUOT
SLM tu- ma- liyat atu aapa
SLMT paa -kk.a ille enna eṇḍu
NEG INF care.for INF NEG one what QUOT
SLM suḍa kitan-na an- liyat si
SLMT ini engaḷ-e paa -pp -aangaḷ aa
LINK 1PL-ACC FUT care.for FUT AGR Q
'Now they don't care for their [own] village folk, such as their elder and younger siblings – not a bit! So will they care for us?'

3.1.4 *Lack of tense marking*

Main verbs are usually tense-marked, but a few examples, such as (18)–(20), show that tense-marking is not obligatory. This is likely a carryover from VM, where time reference is usually inferred from the context. Another such carryover is the behaviour of the existential quasi-verb *aaɖa*, which as illustrated in (20) is never marked for PRESENT. In VM *ada* differs from other verbs in not permitting a full range of pre-verbal markers. Since the Tamil counterpart *iru* inflects like a normal verb, the exceptionality of *aaɖa* is one small indication that SLM is not simply a calque of Tamil.

(18) SLM ini batu-ka joo bisaar trabal ayn- jaaḍi
SLMT inda kallu-le daan periya trabal ʋa -nd -uccu
this rock-LOC FOC big trouble PAST happen come PAST AGR
A: 'Right at this 'rock' [a granite hill] there was big trouble'
SLM batu-ka aapa-aapa jaaḍi
SLMT kallu-le enna-enna naɖa -nd -uccu
rock-LOC what-what happen PAST PAST.AGR
B: 'What happened at the rock?'

(19) SLM lay.le itu seksa-nya kitam baa-pii[9]
SLMT inn.um anda kareccal-e naanga koɳɖup-poo -r -oom
still that trouble-ACC 1PL take-go PRES AGR
'We still experience [lit. carry] that trouble'

(20) SLM itu-pe kopi atu aaɖa si
SLMT adu-ɖa piradi oɳɖu iru -kk -ud aa
that-GEN copy one be PRES AGR Q
'Is there a copy of that?'

3.2 Aspect

In SLM, as in Tamil, PERFECT, PERFECTIVE, and PROGRESSIVE aspects can be marked. (Sinhala also has PERFECT and PROGRESSIVE as grammatical categories; perfective can be marked, but it is not a grammatically integrated category.)

3.2.1 *Perfect: V + aaɖa*

The perfect, formed by the verb followed by *aaɖa* functioning as auxiliary, indicates the relevance of events or states begun (and often completed) prior to the reference time. Example (21) illustrates the formation, which is clearly calqued, as Tamil (and Sinhala) have the same structure. One difference, noted earlier, is that *aaɖa* does not inflect for tense, while the Tamil auxiliary *iru* is marked for tense (and agreement) yielding a present-perfect/pluperfect contrast not found in SLM. Also, the verb is in

9. This is an unusual use of this verb.

participial form in Tamil (and Sinhala). As noted earlier, double past marking can be used to mark participles in SLM, as in (23), but the distinction between participles and main verbs is not regularly maintained. As in Tamil, the SLM perfect may also indicate remote past; this function is illustrated in (25).

(21) SLM braapa ooram -paɖa siitu pii aaɖa
SLMT ettane peer ange poo -y iru -nd -aanga
how many person PL there go PPL PRF PAST AGR
SLM samma.oorang.le siini daatang aaɖa
SLMT ellaar.um inge ʋand ii -kk -aanga
everyone here come PRF PRES AGR
'How many people had gone there?. . ..Everyone has come back here?'

(22) SLM loram-pe baapa aapa pĕkajang kĕja aaɖa
SLMT onga-ɖa ʋaappaa enna ʋeele se -nj -u iru -nd -aar
2PL-GEN father what work do PAST PPL PRF PAST AGR
'What work did your father do?'

(23) SLM banyak musing -ka naasi-muʈʈi sama daapuru-ka[10]
SLMT miccam kaalam -aa soottu-muʈʈi-ooɖe aɖuppu-le
much time LOC ADVZ rice-pot ASSOC
SLM kitang uutang-ka as- jaaga aaɖa
SLMT naanga kaaʈʈu-le paatukaa -tt -u ii -kk -oom
1PL jungle-LOC PAST.PPL safeguard PAST PPL PRF PRES AGR
'For a long time we have hidden in the jungle with [only our] rice pots and [our] charcoal cooker'

(24) SLM deram-pe ʋaralaaru sama ayng- kumpul aaɖa
SLMT avanga-ɖa ʋaralaaru ellaam seekkap.paaʈ -ʈ -u ii
3PL-GEN history all PAST collect PAST PPL PRF
SLM
SLMT -kk -udu
PRES AGR
'Their history etc. has been collected' [Lit. Have collected their history.]

(25) SLM itu-na le blaakang banyak ooram-paɖa
SLMT adu-kk um muɳɖi miccam aaɭ-kaɭ
that-DAT EMPH after many person-PL
SLM ara- niingal aaɖa
SLMT maʋtt.aaʋ -i iru -nd -aanga
PRES die PPL PRF PAST AGR
'Even after that many people [who lived here] had passed away'

10. The use of the locative is odd here.

3.2.2 *Perfective: V + abis; V + taaro (etc.?)*

Although both PERFECT and PERFECTIVE may often be translated by the English perfect, in SLM (and Tamil) they contrast. The PERFECTIVE focuses on the completion of an event rather than its relevance. Its English translation may often employ a phrasal verb including a perfective particle (e.g., *finish off, smash up, break down*). The PERFECTIVE is formed, as in Tamil, by adding a perfective auxiliary to the verb. The unmarked perfective auxiliary is *abis*, illustrated in (26)–(27), deriving from the VM postverbal marker *habis* 'finished', which frequently has a perfective sense. Note that in Tamil the unmarked perfective auxiliary, *viɖu*, usually reduces and encliticizes to the main verb, which is in past participial form, and the auxiliary itself inflects for tense and agreement. In SLM, the auxiliary is unmarked for tense (or agreement); the main verb is optionally marked past.

An alternative perfective auxiliary *taaro* 'put', seen in (28), usually implies willful action with irreversible change of state or condition. The auxiliary *taaro* derives from VM *taro* 'put', which is not used aspectually in either Std. Malay or VM. The SLM usage is clearly calqued on the Tamil equivalent *pooɖu* 'put' (PFV). In Tamil, a number of such alternative perfective auxiliaries are available, each carrying additional semantic information (See Asher 1985: 178ff.; Schiffman 1999: 82ff.). It would not be surprising to find that SLM has additional alternative perfective auxiliaries, although none appears in our data.

(26) SLM kitang suɖa::: itu -na nya- soovek abis
SLMT naanga ini atu-vaɭ -e kiɭi -cc -i -ʈ -ʈ -oom
1PL LINK that-PL -ACC PAST tear PAST PPL PFV PAST AGR
'We::::ll, we have torn them up'

(27) SLM see itu-na juval abis
SLMT naan ad-e vit -t -i -ʈ -ʈ -an
1sg that-ACC PAST sell PAST PPL PFV PAST AGR
'I sold it off' (elicited)

(28) SLM ziyaaram-na suɖa picakan taaro
SLMT ziyaratt-e oɖe -cc -i pooʈ -ʈ -aanga
shrine-ACC PAST break PAST PPL PFV PAST AGR
'They tore down the shrine'

3.2.3 *Progressive: (ayn-/su-)V + ambel + aaɖa*

The PROGRESSIVE formation, illustrated in (29)–(30), is clearly calqued on the corresponding Tamil formation V-PAST-PPL + *koɳɖu* + *iru*. The parallels between the two structures are threefold. First, both use a past form of the verb: in Tamil the participle is built from the PAST stem; past marking is optional in SLM. Second, the progressive auxiliary *ambel/koɳɖu* has two other functions in both languages: a main verb meaning 'take' and a marker of the so-called reflexive/self-benefactive. Finally, both constructions use the existential verb *aaɖa/iru* as a second auxiliary in main clauses. VM marks

progressive with *lagi* or *ada; ambil* occurs only as a main verb meaning 'take', 'grasp'. Std. Malay has an unrelated progressive marker *sĕdang*.

(29) SLM deraŋ daataŋ dupaŋ n- duuɖuk
SLMT aʋanga ʋar-adu-kku muɳɖi iru -ndu
3PL come-NOMZ-DAT before PAST be PAST
SLM siini islaam ayn- duuɖuk ambel joo aaɖa, tuʋan
SLMT inge islaam iru -nd -u koɳɖu ttaan iru -ndu, tura
here Islam PAST be PAST -PPL PROG FOC be PAST sir
'[Islam] was [here] before they came, Islam was [present] here [all along], sir' [Lit: was being here/ was staying here]

(30) SLM kaaram suɖa kubaali le ini reepot atu
SLMT ippa ini marappayinam um inda karaccal oɳɖu
now LINK again EMPH this trouble one
SLM su- baaʋung ambel aaɖa
SLMT ɛlumb -i koɳɖu iru -kku
PAST arise PPL PROG be PRES
'Now, see, this trouble is arising again'

3.3 Mood

SLM mood markers parallel those of Tamil. The scope of this paper does not permit a complete account of mood and modality in the two languages, but we survey selected mood markers. The classificatory framework adopted is that of Palmer (2001), outlined in Table 2.

Table 2. Palmer's (2001) classification of modality

Propositional Modality: factual status		Event Modality: person-event relationship	
Epistemic: judgements speculative, deductive, assumptive	Evidential: evidence sensory, reported	Deontic: External obligation, permission, commitment	Dynamic: internal volition, ability

3.3.1 *Event:dynamic:volitive:* N-DAT . . . (*ma-*)*V*(*na-*) + *maaʋ*

The VOLITIVE construction, illustrated in (31)–(32), is calqued on the Tamil structure V-INF + *ooɳum*. In our SLM data the verb is unmarked, but P. Slomanson has observed variable marking of non-finite status (p.c., October 2005); in both languages the construction takes a dative subject, as seen in (31). The corresponding formation in VM is *mau* + V with a nominative subject.

(31) SLM ini lorampaɖa-na siini-ka duuɖuk maaʋ si, lay
SLMT idu ongaɭ-ukku inge iru -kk.a ooɳum aa, innum
So 2PL-DAT here LOC be INF VOL Q still
'So, do you want to stay here? – still?'

(32) SLM aʋtsteeshan-na pii maaʋ
SLMT ʋeɭiyuur-ukku poo -g.a ooɳum
oustation-DAT go INF VOL
'[We] want to go to an outstation'

3.3.2 *Event:deontic:debitive:* N-NOM (*masi-/ma-/mang-*)*V* + *maaʋ*
The DEBITIVE construction, illustrated in (33)–(36), is similar in form to the volitive, but its subject is nominative rather than dative, and the verb usually bears a prefix, though bare verbs also occur, as in (36). The multiple forms of the prefix require discussion. First, note that the VM preverbal DEBITIVE marker *masi/mesi* (Std. Malay *mĕsti*), is clearly the source of the SLM prefix *masi-* seen in (33). It may also be the source, via reduction, of the prefix *ma-* seen in (34).[11] The third prefix *mang-*, seen in (33) and (35) appears to derive formally from the Malay verbal prefix *mĕN-*, but its function in SLM has been analyzed as an infinitive marker (Saldin 1996; Hussainmiya 1986) or as a marker of non-finiteness on nominalized verbs (Slomanson 2005, 2006). Since it also has a reduced allomorph *ma-*, its occurrence in debitive constructions may result from hypercorrection. In Tamil the debitive construction is V-INF + *ooɳum*, which is identical to the volitive construction except that, as in SLM, its subject is nominative rather than dative. Since *mau* may not cooccur with *mes*(*t*)*i* with this meaning in Vehicular or Std. Malay, its appearance in the SLM debitive is clearly due to Tamil influence. When the subject is omitted, the Tamil construction can only be distinguished by context from the volitive; the same is true in SLM if the prefix *ma-* is used.

(33) SLM siini-ka jo man- duuduk maaʋ lorampada sama
SLMT inge taan iru -kk.a ooɳum niinga ellaar.um
here LOC FOC INF live INF DEB/VOL 2PL all
SLM masi- pii maaʋ kata kitan-na asi- puukul
SLMT poo -g.a ooɳum eɳɖu engaɭ-a aɖi -cc -i
DEB go INF DEB QUOT 1PL-ACC PAST.PPL beat PAST PPL
SLM kitam-pe ruuma-pada-na sĕ- baaka
SLMT enga-ɖa uuɖu-ʋaɭ-a paʈʈa.ʋe -cc -aanga
1PL-GEN house-PL-ACC PAST burn PAST AGR
'Saying 'we have/want to live right here; you all must go,' they beat us and burnt our houses'

11. P. Slomanson (p.c., October 2005) argues that "mau is a lot more convincing as an etymon for ma- because of the want→future→non-finite path, which is not unique to this language."

(34) SLM derampaɖa-na laas-na [DAT] jop ma- kaasi maaʋ
SLMT aʋangaɭ-ukku kadeesi-la [LOC] jop-ʋaɭ-a kuɖu -kk.a ooɳum
3PL-DAT end-CASE job-PL-ACC INF give INF DEB
'In the end they had to be given jobs' [Lit. In the end had to give them jobs]

(35) SLM ini kaaram derang-ka jo kitang klaapa
SLMT indaa ippa aʋangaɭ-ukkiʈʈa taan naanga teengaa
here now 3PL-(COL)LOC FOC 1PL coconut
SLM mam- biili maaʋ běras mam- biili maaʋ
SLMT ʋeeng -a ooɳum arasi ʋeeng -a ooɳum
INF buy INF DEB rice INF buy INF DEB
SLM kitam-pe ikan-nang deran-na man- juuʋal maaʋ
SLMT enga-ɖa miin-a aʋangaɭ-ukku ʋi -kk.a ooɳum
1PL-GEN fish-ACC 3PL-DAT INF sell INF DEB
'Here, now it's from them that we must buy our coconuts, [and] must buy rice; we have to sell our fish to them'

(36) SLM kitan-na aayir tira kalo
SLMT engaɭ-ukku taɳɳi ill -aaʈʈi
1PL-DAT water lack COND
SLM deram-pe nigrii-na pii maaʋ, mandi maaʋ
SLMT aʋanga-ɖa uur-ukku poo -g.a ooɳum, kuɭi -kk.a ooɳum
3PL-GEN village-DAT go INF DEB bathe INF DEB
'If we lack water, we must go to their village, [and] we must bathe [there]'

3.3.3 *Event:dynamic:habilitative: (mang-)V-na + boole*

The habilitative construction is illustrated in (37) and (38). The formation is calqued on the Tamil structure V-INF + *eelum*. In SLM the verb bears the infinitival suffix *-na* which is frequently accompanied by the prefix *mang-* and its variants, which we have also glossed as an infinitive marker. The corresponding formation in Malaysian VM is *boleh* + V and in Indonesian VM *bisa* + V; thus the SLM formation represents the adoption of Tamil word order and the requirement for infinitival marking on the verb.

(37) SLM ciŋggala-paɖa-na le baay-na mlaayu oomon -na boole
SLMT singaɭaʋan-gaɭ-ukk um nall-aa jaaʋam pees -a eelum
Sinhalese-PL-DAT CONJ good-ADVZ Malay speak INF HABIL
'The Sinhalese also can speak Malay well'

(38) SLM kitam-pe aasal-ka joo kita ma- pii -na boole
SLMT enga-ɖa sondatt-ukkiʈʈa taan naanga poo -g.a eelum
1PL-GEN origin-(COL)LOC FOC 1PL INF go INF HABIL
'It is to our original home that we can go'

3.3.4 *Propositional:evidential:reportative: kanyang*

The reportative, illustrated in (39)–(40), is a clause-final particle indicating that the speaker assigns responsibility for the truth of the proposition to a third party. Analogous clause-final reportative markers are found in both Sinhala (*-lu*), Tamil (*aam*) and Sri Lankan English (*it seems*). The source of *kanyang*, which appears as *kiyang* in Bichsel-Stettler's texts (1989), is uncertain; likely candidates are the reportatives *kayanya* (Jakarta Malay) and *katanya* (VM) via nasal spreading and reduction. Std. Malay also has a clause initial reportative *rupanya.*

(39)	SLM	ini	kumbali	tĕra	kaasi			kanyang
	SLMT	itu	marappaḍiy.um		kuḍu	-kk.a	illay	aam
		this	again	NEG	give	INF	NEG	REP
		'They won't give this again, it seems' (Sri Lankan Eng.)						

(40)	SLM	itu	muusing-ka	aaḍa			kanyang
	SLMT	anda	kaalatt-ila	iru	-nd	-ucc	aam
		that	time-LOC	be	PAST	AGR	REP
		'There are said to have been [some Jawi-script books] in those days'					

4. Discussion

4.1 Origin of TMA in Sri Lanka Malay

TMA has long been a focus of creole studies, and especially since the pioneering work of Derek Bickerton (e.g. 1980 [1974] and 1981) the categories ANTERIOR, NONPUNCTUAL, and IRREALIS, together with their unmarked counterparts, have been associated with creoles. Although SLM differs in its sociolinguistic history from the creoles on which these archetypal features have been established, and although the universality of Bickerton's system is in doubt (e.g., see discussion in Holm 2000: 174–194), we must at least entertain the possibility that a source of TMA in SLM can be found in universal properties of the creolization process. It should be clear from the foregoing, however, that the TMA categories of SLM more closely match those of Sri Lanka Muslim Tamil than those of the archetypal creole system or any characterization under the rubric of UG or the 'bioprogram'. In many instances the SLM expressions are clearly calqued on the corresponding Tamil expressions; in other cases the categories match, but the morphosyntactic realization of these categories differs. Traces of the VM TMA system can also be discerned, particularly in the pre-verbal position of some of the markers.

Before concluding that Tamil has influenced TMA marking in SLM, it is necessary to show that at least some of its Tamil-like properties are not derived from VM. The following points can be made:

a. SLM and Tamil mark three tense categories, PAST, PRESENT and FUTURE, while VM marks none. Occasional lack of marking in SLM may be a reflection of the

VM system. The SLM future marker is derived from a VM adverbial; the present and past tense markers are reinterpreted VM aspectual markers.

b. Both SLM and VM mark PERFECTIVE and PROGRESSIVE; however the SLM marking is calqued on Tamil and unrelated to that of VM. In addition, SLM and Tamil mark PERFECT, a category not found in VM. SLM has simply abandoned the VM aspect marking system and built its own on the basis of Tamil.

c. In mood more than the other two systems, SLM displays its mixed heritage, perhaps because the categories marked are found in both Tamil and VM. The mood markers of VM are found in SLM, but their syntax is again partly or fully calqued on Tamil.

We conclude that, to a large extent, the TMA system in SLM has been created through reinterpretation of existing VM markers on the basis of Tamil categories and through calquing of Tamil expressions. These same processes have been noted by Smith (2001) in other creolization and contact situations. Why does TMA in SLM not resemble the creole archetype? The main reason is that the creole archetype is biased toward Atlantic creoles and reflects their West-African substrates, rather than the typologically Dravidian substrate of SLM. It is also notable that the social context of creolization for SLM was different from that of archetypal creoles. In particular the substrate, being homogeneous, could have been expected to have a stronger effect than the heterogeneous substrate assumed of archetypal creoles. We have assumed that the principal substrate language was Tamil, but even if Sinhala were admitted to have had an influence, the fact that its TMA categories are similar to those of Tamil means that the two languages would have acted in concert rather than in competition.

4.2 Timing of the development of Tamilized TMA in SLM

We now turn to the question of the timing of the development of Tamilized TMA in SLM. In particular we address the issue of whether this is a recent development, as suggested by Bakker (2000a, 2000b), or whether it derives from the creolization situation itself, as we maintain. We begin with a discussion of the socio-historical situation.

4.2.1 *Socio-historical situation*

The social history of the SLM community had an important bearing on its linguistic history and on the development of SLM itself. The important events for our discussion are found in Table 3.

From their earliest days in Sri Lanka, 'Malays' had a close relationship with the Tamil-speaking 'Moors' (see 1.1). Intermarriage was first noted in 1680 by a German observer, Christoph Schweizer [a.k.a. Christopher Schwitser], who also commented on the multilingualism of the Malay soldiers (Hussaimiya 1987: 49–50), and was confirmed by Dutch *thombos* (administrative records) of the eighteenth century (Hussainmiya 1987: 52). An indigenized variety of Malay must have arisen among the children of these marriages, whose fathers spoke VM but may have been familiar with

Table 3. Milestones in the development of Sri Lanka Malay (dates from Hussainmiya 1987)

Date		Event
1600		
	1640	Dutch attack Portuguese in Galle, assisted by soldiers from the East Indies.
1650		
	1656	Dutch gain control of coastal Ceylon from Portuguese. Soldiers, convicts and other political exiles are brought from the East Indies for the next 140 years.
	1680	First written reports of intermarriage between East Indians and Sri Lanka Moors. First reports of use of the Malay language in Sri Lanka.
1700		
1750	1760s	Dutch administrative records (*Thombos*) report intermarriage between Malays and Moors.
	1764	2,408 Malays reported serving in Dutch army in Sri Lanka.
	1796	British gain control of coastal Sri Lanka from the Dutch, begin waves of migration from Malaya and establish military schools beginning around 1802 in which Malay language is taught.
1800	1811–16	British rule in East Indies. Renewed migration from Java and Madura.
	1820	Date of oldest surviving original text from Sri Lanka: Syair Kisahnya Khabar Wolenter Orang Benggali.
1850	1869–78	Alamat Langkapuri newspaper published: first Malay newspaper anywhere in Jawi script. Traditional Malay literature widely read and copied.
	1873	Ceylon Rifle Regiment disbanded.
	1895–99	Wajah Selong newspaper published; distributed in Malaya and Dutch East Indies.
1900	1906	Two collections of Malay pantuns published, both in Roman script with influence from colloquial SLM.
	1930s	Last recorded copying of traditional Malay literature in Sri Lanka.
	1940s	Last known Malay language correspondence in Sri Lanka.
1950	1970s	First linguistic studies of SLM; first recorded evidence of colloquial SLM.
2000		

Tamil through their interactions with the Moor community, and whose mothers spoke Tamil natively and must have attempted to learn Malay without the benefit of instruction. Tamil-speaking mothers must also have tried to make Malay the language of the home; otherwise it would be difficult to imagine the language being maintained by subsequent generations.

There is no evidence that Malay served as a literary language during this early period. VM was a strictly colloquial language and, barring a few minor exceptions, not written anywhere in the Malay world, this function being served by Std. Malay. Soldiers, recruited from various parts of the East Indian archipelago, were unlikely to have been literate in Malay, and no writing in Malay has been found from this period.

Soon after the British took control of Sri Lanka in 1796, Malay-language education was established for soldiers and their families (Hussainmiya 1987: 93). The variety taught was standard literary Malay, which differed significantly from all colloquial

varieties. The resultant diglossia was consistent with the pattern already established throughout the Malay-speaking world and also found in Tamil, Sinhala and Sri Lanka Portuguese (Smith 1998). The diglossic high was Std. Malay, written in *Jawi* (modified Arabic) script, and the low was SLM.

There was an extensive literature in Malay written in Sri Lanka through much of the nineteenth century, consisting of copies of traditional epics and religious tracts, with some works originally written in Sri Lanka (Hussainmiya 1987; Harun 1986). In addition, two Malay-language newspapers appeared. In all cases, the written language was standard literary Malay, with only minor influence from the colloquial language.

In 1873, the British disbanded the CRR, consequently halting education in the Malay language; a precipitous decline in Malay-language literacy ensued. In the early years of the twentieth century, several works including collections of Malay *pantuns* (popular poetry), appeared. These collections were written in Roman script and show influence from colloquial SLM (Hussainmiya 1987: 97; Harun 1986: 75). They were, however, isolated occurrences, and no further literature in high or low Malay subsequently appeared. By the mid-twentieth century, literacy in any form of Malay in Sri Lanka was completely lost.

Within the framework of the development of SLM over the past 350 years, the rise and fall of literacy in High Malay stands as a relatively brief interlude of 70 years, with little if any effect on the development of the colloquial language.

4.2.2 *Conversion or creolization?*

In this context, let us examine Bakker's claim that SLM is a 'converted' language (2000a: 33), a language which has undergone massive "grammatical influence without lexical influence" (Bakker 2000b: 607). Bakker (2000a) argues that SLM has "been Tamilized in a very short time span" (2000a: 33),[12] claiming that in 19th C documents, SLM has a TMA system "which can be considered typical for creole languages" (2000a: 31).[13] We have pointed out that the 19th century documents are in High Malay; they tell us nothing about the spoken SLM of the time and therefore do not support Bakker's claim of recent rapid Tamilization. Indeed, two facts indirectly indicate that the 19th century spoken language was already quite different from the literary: first the fact that literacy in SLM declined so rapidly after the cessation of formal instruction; second the fact that barely two decades later, documents exhibiting features of collo-

12. Bakker makes the same claim for Sri Lanka Portuguese; see Smith (1979a, 1993, 1998) for counterarguments.

13. Bakker (2000b) adds that "earlier documents show a more ordinary form of Malay" (2000b: 607). Taken together, the two claims seem to imply that Malay looks like a typical creole in its TMA and other properties. While some have claimed that VM has been pidginized (e.g. Adelaar & Prentice 1996: 674), neither pidginization nor creolization are evident in the 19th Century SLM documents, but these, as noted, are in Std. Malay.

quial SLM appear. As Smith (2003) points out, a third factor that mitigates against the likelihood of rapid contact-induced change at the end of the 19th century is that

> such change would be expected only in the context of radical social change. The only major social event in the late 19th century is the disbandment of the Ceylon Rifle Regiment. However wrenching this may have been for the lives of individual Sri Lanka Malays, it can hardly be considered the type of revolutionary upheaval required for rapid contact-induced change. For that, we need to look at the creolization process itself, when Malay-speaking men married local, mainly Tamil-speaking women and a new local variety of Malay was created within these mixed-marriage families.

While it is certainly possible that a few individual Tamilized features of SLM may have developed through later areal convergence, the most likely source of the majority of such features is the initial contact situation.

5. Summary

We have argued that Tamil is the main source of the TMA categories of SLM, while their formal expression results from a mixture of reinterpretation of VM forms and calquing of Tamil expressions. SLM TMA closely matches that of Tamil and differs substantially from the systems found in both Standard and Vehicular Malay as well as from the system often associated with archetypal creoles. We have argued that the similarity to Tamil results from the creolization process rather than from more recent convergence. We have also argued that the difference between SLM TMA and that of the creole archetype is due largely to the fact that SLM had a homogeneous Tamil (or at least, typologically Dravidian) substrate.

References

Adelaar, K. A. (1991). Some notes on the origin of Sri Lanka Malay. In H. Steinhauer (Ed.), *Papers in Austronesian Linguistics*, No. 1 (pp. 23–37). Canberra: Research School of Pacific Studies.

Adelaar, K. A. & Prentice, D. J. (1996). Malay: Its history, role and spread. In S. A. Wurm, P. Mühlhäusler, & D. T. Tryon (Eds.), *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas* (Volume II.1: Texts) (pp. 673–693). Berlin: Mouton de Gruyter.

Alwi, H., Dardjowidjojo, S., Lapoliwa, H., & Moeliono, A. M. (1998). *Tata Bahasa Baku Bahasa Indonesia*. Jakarta: Balai Pustaka.

Asher, R. E. (1985). *Tamil* [Croom Helm Descriptive Grammars]. London: Croom Helm.

Bakker, P. (2000a). Convergence intertwining: An alternative way towards the genesis of mixed languages. In D. Gilbers, J. Nerbonne, & J. Schaeken (Eds.), *Languages in Contact* (pp. 29–35). Amsterdam: Rodopi.

Bakker, P. (2000b). Rapid language change: Creolization, intertwining, convergence. In C. Renfrew, A. McMahon, & L. Trask (Eds.), *Time Depth in Historical Linguistics* (Volume 2) (pp. 585–619). Cambridge: The McDonald Institute for Archaeological Research.

Bichsel-Stettler, A. (1989). Aspects of the Sri Lanka Malay Community and its Language. MA thesis, University of Bern.

Bickerton, D. (1980 [1974]). Creolization, linguistic universals, natural semantax and the brain. In R. R. Day (Ed.), *Issues in English Creoles: Papers from the 1975 Hawaii conference.* Heidelberg: Groos.

Bickerton, D. (1981). *Roots of Language.* Ann Arbor MI: Karoma.

Gair, J. W. (1976). The verb in Sinhala, with some preliminary remarks on Dravidianization, *International Journal of Dravidian Linguistics, 5*, 259–273.

Harun Mat Piah (1986). Tradisi kesusasteraan Melayu Sri Lanka dalam konteks kesusasteraan Melayu tradisional nusantara: Satu tinjauan ringkas. *Sari, 4*(2), 63–82.

Holm, J. (2000). *An Introduction to Pidgins and Creoles.* Cambridge: CUP.

Hussainmiya, B. A. (1986). 'Melayu Bahasa': Some preliminary observations on the Malay creole of Sri Lanka. *Sari, 4*(1), 19–30.

Hussainmiya, B. A. (1987). *Lost Cousins: The Malays of Sri Lanka*. Bangi: Penerbit Universiti Kebangsaan Malaysia.

Hussainmiya, B. A. (1990). *Orang Regimen: The Malays of the Ceylon Rifle Regiment.* Bangi: Penerbit Universiti Kebangsaan Malaysia.

van Minde, D. (1997). *Malayu Ambong: Phonology, morphology, syntax.* Leiden: CNWS.

Masica, C. P. (1976). *Defining a Linguistic Area: South Asia.* Chicago IL: University of Chicago Press.

Paauw, S. (2003). What is Bazaar Malay? Paper presented at the Seventh International Symposium on Malay/Indonesian Linguistics, Nijmegen, The Netherlands.

Paauw, S. (2004). *A Historical Analysis of the Lexical Sources of Sri Lanka Malay.* MA thesis, York University.

Paauw, S. (2005). Malay dialectology: A new analysis. Paper presented at the Niagara Linguistic Society, Buffalo, New York.

Palmer, F. R. (2001). *Mood and Modality* (2nd edition). Cambridge: CUP.

Robuchon, G. (2003). Malayo language in Sri Lanka. Paper presented at the Seventh International Symposium on Malay/Indonesian Linguistics, Nijmegen, The Netherlands.

Saldin, B. D. K. (1996). *The Sri Lankan Malays and Their Language* (*Orang Melayu Sri Lanka dan bahasanya*). Dehiwala: Sridevi Printers Publication.

Schiffman, H. F. (1999). *A Reference Grammar of Spoken Tamil.* Cambridge: CUP.

Slomanson, P. (2005). The Verbal Morphosyntax of Non-Canonical Contact Languages – Malay-derived Constraints and the Inflectional Domain in Afrikaans and Sri Lankan Malay. Ph.D dissertation, City University of New York Graduate Center, New York.

Slomanson, P. (2006). Sri Lankan Malay Grammars: Lankan or Malay? This volume.

Smith, I. R. (1978). *Sri Lanka Portuguese Phonology*. Trivandrum: Dravidian Linguistics Association.

Smith, I. R. (1979a). Convergence in South Asia: A creole example. *Lingua, 48*, 193–222.

Smith, I. R. (1979b). Substrata versus universals in the formation of Sri Lanka Portuguese. *Papers in Pidgin and Creole Linguistics* No. 2 [Pacific Linguistics Series A No. 57], 183–200. Canberra: Australian National University.

Smith, I. R. (1993). Review of Jackson, K. David. *Sing without Shame: Oral traditions in Indo-Portuguese verse with transcription and analysis of a Nineteenth-Century Manuscript of Ceylon Portuguese*. (John Benjamins, 1990). *Journal of Pidgin and Creole Languages, 8*, 291–296.

Smith, I. R. (1998). Introdução. *Dialecto Indo-Português de Ceylão*, 11–37. Lisbon: Comissão Nacional para as Comemorações dos Descobrimentos Portuguêses.

Smith, I. R. (2001). Dravidian negatives in Sourashtra and Sri Lanka Portuguese. Paper presented at the Linguistic Society of America annual meeting, Washington DC.

Smith, I. R. (2003). The provenance and timing of substrate influence in Sri Lanka Malay. Paper presented at the South Asian Language Analysis Roundtable XXIII, Austin, Texas.

Smith, I. R., Paauw, S., & Hussainmiya, B. A. (2004). Sri Lanka Malay: The state of the art. In R. Singh (Ed.), *The Yearbook of South Asian Languages and Linguistics 2004* (pp. 197–215). Berlin: Mouton de Gruyter.

Sneddon, J. N. (1996). *Indonesian Reference Grammar*. St. Leonards NSW: Allen & Unwin.

[Sri Lanka 2001 Census] Sri Lanka Department of Census and Statistics. 2003.10.4. *Census of population and housing, 2001. Population by Ethnicity according to District and Sector* (*Provisional*). Retrieved 2003.10.07 from http://www.statistics.gov.lk/census2001/population/district/t001c.htm

The advantages of a blockage-based etymological dictionary for proven or putative relexified languages

(Extrapolating from the Yiddish experience)

Paul Wexler
Tel Aviv University, Israel

Yiddish is generally regarded as a form of High German that underwent extreme Slavicization over some seven centuries. I believe, in contrast, that Yiddish is a mixed Slavic language that was created by two separate processes of relexification between the 9th–15th centuries: from Upper Sorbian and Kiev-Polessian (contemporary North Ukrainian and Southern Belarusian) to Middle High German and, secondarily, Classical Hebrew. The relexification was described in great detail in earlier studies of mine. The present paper shows (1) why an etymological dictionary of a relexified language must be radically different from that of a non-relexified language and (2) why such a dictionary is an ideal way to demonstrate the fact of relexification.

1. Introduction: An overview of relexification in the history of Yiddish

Back in 1991 I proposed that Yiddish was not a Germanic language but rather a West Slavic language whose vocabulary had been largely relexified to High German; Hebrew was a secondary lexifier source. I speculated that Jews speaking Sorbian between the 9th–12th centuries had resisted the pressure to switch to German and Christianity as Germans migrated into the mixed Germano-Slavic lands. While most of the Slavs eventually became Germanized, the Jews – mainly a local proselyte population descended from a minority ethnic Palestinian Jewish population – made only a partial language shift to German that entailed keeping Sorbian grammar, phonology and phonotactics, while replacing most of the Sorbian lexicon with German phonetic strings which received their lexico-semantic and syntactic parameters exclusively from the Slavic substratum; this process is known as relexification (see Lefebvre 1986; Horvath & Wexler 1994, 1997).

The purpose of the present paper is to suggest a new format for an etymological dictionary of a proven or putative relexified language; the language of illustration is

Yiddish, a language I believe was created by relexification. Some motivation for this claim is presented here in order to motivate the lexicographical proposals, but for detailed discussion, the reader is directed to Wexler 2002b.

In recent years it has become clear that the 1991 model of a single Sorbian substratum is insufficient to account for the genesis of Yiddish; specifically, other developments in Eastern Europe had immediate bearing on the topic. For example, many Slavic(ized) descendants of the Khazar Jews settled both before and after the collapse of their Khaganate in the 10th century in the Kiev-Polessian principality (southern Belarus and northern Ukraine). At the time that the Sorbian Jews first brought Yiddish into the East Slavic domain in the 13th century, the indigenous Jews were still speaking Kiev-Polessian (the original common ancestor of North Ukrainian and South Belarusian dialects). A second relexification act must have taken place when Kiev-Polessian speakers adopted Yiddish (and new High German) vocabulary, while continuing to use their grammar; the second relexification phase might even have begun in Central Europe in the 9th century, if not earlier, when Judaized Khazars began to settle there. Given the close similarities between Old Sorbian and Kiev-Polessian grammar, few radical changes would have been expected in Yiddish grammar as a result of the second relexification process. Contemporary Yiddish can thus be described as the combined product of two relexification processes from West and East Slavic. I conjecture that the first Jews to reach the Sorbian lands were mainly male immigrants from the Balkans, accompanying Serbs in their migration north to the German lands after the 6–7th centuries; some of these migrants most likely spoke a South Slavic language (Wexler 1992). Hence, some Yiddish patterns of discourse that can in theory be ascribed to Kiev-Polessian or to Upper Sorbian may ultimately have their origins in South Slavic. Significantly, unrelexified relic East Slavisms in Yiddish are very often characteristic of South and West Belarusian and North and West Ukrainian – precisely the areas of the original Kiev-Polessian dialect up until its disintegration around 1400 and a major domicile of Judaized Khazars.

The Sorbs, Avars, Khazars and others who converted to Judaism in the first millennium were seeking a new ethno-religious identity – a prerequisite for relexification, in the first phase, to avoid enslavement by the Germans, and in the second phase, to have a neutral religion vis-à-vis Christian Byzantium and the Muslim Caliphate in Iraq (see Section 2 below). Hence, it was originally non-Jews who were the major players on the stage of the 'Ashkenazic' ethnogenesis and the initiators of the relexification processes and not the small Palestinian Jewish founder population (Wexler 1991, 2001, 2002a, 2002b).

2. Identifying relexification processes and lexical blockage

I would propose nine diagnostic tests for the identification of prior relexification as the driving force behind the genesis of a language, of which the first three are major; the others are, at best, suggestive (e.g. conditions #4 and #7 can obtain in the absence

of relexification). The two central tests for relexification, ##1–2 outlined below, should be carried out simultaneously.

1. Yiddish is a relexification-based language because its grammar and phonology are derived predominantly from a single source, Slavic, while the bulk of the lexicon has a different source (c. 75% is of High German origin). The remaining lexicon is split roughly between Hebraisms (and innovative pseudo-Hebraisms or 'Hebroidisms') and Slavisms – the sole native elements. These claims are supported by a steadily growing body of literature contrasting German, Slavic and Yiddish phonology, grammar and lexicon (e.g. Wolf 1956; Lötzsch 1990, 1996; Geller 2001). Unfortunately, appreciation of these facts (e.g. a derivational morphology calibrated according to Slavic parameters, a Slavic phonology and phonotactics, Slavic word order) is often complicated since (a) many contemporary dialects of Yiddish have been Germanized in the last two centuries (especially in areas where Yiddish speakers were also fluent in German, such as Latvian Kurland and Ukrainian Bukovina), and (b) most Yiddish texts before the early 19th century were written in a Germanized Yiddish – even by speakers of Slavic Yiddish.

2. A language is relexification-based if, by comparing the lexicons and derivational structures of the putative substratal and superstratal (lexifier) languages, it is possible to *predict* with high accuracy (i.e. to motivate *ex post facto*) which superstratal lexical elements will be compatible with the substratal grammar. We can predict accurately which German components would be accepted by Yiddish and where in the lexicon Yiddish will have to use Hebraisms, invent pseudo-Hebraisms ('Hebroidisms' – far more frequent than in any other non-Jewish language adopted and adapted by Jews) and retain unrelexified Slavisms. No other model of Yiddish genesis can make predictions about the component structure of the language. The traditional model of Yiddish genesis that still dominates linguistic and lay circles (see, e.g., the detailed exposition by Vajnrajx 1973) sees the cradle of Yiddish in the Rhineland, when Romance-speaking Jews began to shift to local German dialects; Slavicization of Yiddish allegedly commenced several centuries subsequently. However, this model cannot predict the component structure of Yiddish, or motivate the unique need for artificially created Hebroidisms. Recent attempts to move the cradle of a 'Germanic' Yiddish to Bavaria also have no explanatory power (see e.g. Manaster Ramer 1996).

This second test is extremely important because it shows that it could not have been the seven to eight centuries of contact with Slavic languages that allegedly facilitated German component attrition (in fact, Yiddish was never removed from contact with German in Central and Eastern Europe) or caused extreme Slavicization. Moreover, putative Slavicization of an allegedly 'Germanic' Yiddish could never explain the broad recourse to Hebraisms. If the contacts between Yiddish and the Slavic languages had been a straightforward case of bilingual interference, then the Slavic component in Yiddish should have been (a) found on all levels of the language and (b) entirely haphazard and unpredictable; neither prediction turns out to be the case. It should

also be pointed out that the Slavic 'imprint' in Yiddish differs both quantitatively and qualitatively from that of other languages in contact with Slavic for seven to eight centuries which did not experience relexification, such as Colonial German (see Weinreich 1958) and Karaite in Belarus, Ukraine and Poland (see Wexler 1983).

The Slavic constraints on the acquisition of German result in the blockage of a considerable part of the basic German vocabulary, such that the total German corpus of Yiddish is considerably smaller than the native corpus of any German dialect (e.g. Yiddish practically has no synonyms both of which are of German origin). Blockage manifests itself in the rejection of massive amounts of German vocabulary and morphophonemic alternations that lack Slavic equivalence. The detailed nature of blockage is as follows:

(i) German morphophonemic alternations and derivational morphology are blocked in Yiddish unless they enjoy Slavic parallels. German derivational sets rarely surface in Yiddish *in toto*. The few examples of German Ablaut relationships that are accepted by Yiddish are invariably matched by a Slavic pair, e.g. German *Schweiß* 'sweat'/*schwitzen* 'to sweat' > Yiddish *švejs*/*švicn*, since Slavic also has a common root with alternations, e.g. Upper Sorbian *pót*/*so poćić*, respectively.
(ii) German roots are blocked in Yiddish if Slavic translation equivalents do not broadly overlap semantically.
(iii) Yiddish tends to select German synonyms in accordance with the lexical inventory of Slavic.

Probably all speakers of German today recognize that *schwemmen* 'wash (ashore)' and *schwimmen* 'to swim; float' are related, but probably not that *Ahne* 'ancestor' and *Enkel* 'grandchild' are, though a millennium ago, speakers of German would probably have appreciated the genetic links between the two. Of these examples, Yiddish takes only *švimen* and *enkl*. The reason for the blockage of German *Ahne* is that Slavic has different roots for the set, see e.g. Upper Sorbian *prjedownik, wótc* 'ancestor'/*wnu*(*č*)*k* 'grandchild'. In the first set, Upper Sorbian *pławić* 'to cause to swim, float' is a causative < *płuwać* 'to swim', but the meanings do not fully match those of German *schwemmen/schwimmen.* Partial blockage may also develop because the causative paradigm had become unproductive in one or both of the Slavic substrata, so that speakers sensed no link between the two verbs. It is also possible that Yiddish originally did initially relexify to German *schwemmen*, but lost the term later (in favor of [*lozn*] *švimen*, literally 'let swim') since Ukrainian, the substratum language in the second relexification phase, expresses the two notions of 'float' and 'swim' by different but related roots, e.g. *plavaty, plyvty* 'to float; swim', *vyplyvaty* 'to flow from; swim out; set sail', *pereplyvaty* 'swim across'.

A major clue to the relexifiers' sensitivity to German derivational processes comes from the reliance on Hebraisms, e.g. German *widerspenstig* 'obstinate' and *Gespenst* 'ghost' are historically related to *Gespinst* 'spun yarn, web' and *spinnen* 'to spin'. Yiddish has accepted only the latter two, *gešpin*(*s*) 'spin', *špinen* (*zix*) 'to spin'. The fact that 'obstinate' and 'ghost' are frequently expressed in Yiddish by Hebraisms suggests that the

Sorbian Jewish relexifiers *were* aware of the historical links in German, and refrained from relexifying completely to German since Sorbian used unrelated roots to denote the meanings of the four German terms. This necessitated recourse to Hebraisms.

(iv) German roots which match roots with similar form and meaning in Slavic (the words in question may or may not be cognates) were usually blocked in Yiddish, since they were apparently perceived as Slavicisms. It is difficult to determine how much similarity in form and meaning was required to cause blockage of such Germanisms. German words rejected in relexification could be replaced in four ways: by (a) acceptable Germanisms; (b) Hebrew words; (c) where Hebrew terms are lacking, Yiddish creates Hebroidisms or ascribes new meanings to existing Hebrew words and (d) by retaining substratal Slavisms (and occasionally by creating Slavoidisms). For example, in place of the blocked German *abtrünnig* 'disloyal, unfaithful; apostate', Yiddish uses another Germanism with innovative meanings, e.g. *felčer* 'apostate', *felčn* 'betray' < German *falsch* 'wrong, false', *Fälscher* 'forger', *fälschen* 'forge' (also > Yiddish *falš, felšn*). Thus, the distribution and meanings of German roots in Yiddish often differ substantially from that of the German source etyma. German *Witwe*(*r*) 'widow(er)' would be expected to be blocked in Yiddish because of the existence of cognate Upper Sorbian *wudowa* f, *wudowc* m. Thus Yiddish has *almone* f < Hebrew, and recalibrates *almen* m (< 'lonely man' in Biblical Hebrew).

3. Because of the severe block on German, the primary lexifier source, Yiddish needs recourse to a second minor lexifier, Hebrew/Hebroidisms. No other Jewish language has anywhere near as large a corpus of Hebroidisms as Yiddish. The claim, often made (see e.g. M. Vajnrajx 1973 and Katz 1985), that Hebrew elements were available to speakers of all Jewish languages prior to immigration into the German-speaking lands is preposterous, since there is no evidence of the use of Hebrew in the Khazar kingdom or southern Italy until the late 10th century with Iraqi Jewish westward migrations.

4. We can suspect that Yiddish is relexification-based since semantic domains like Jewish culture and religion tend to be expressed by Slavic elements. These culture-bound terms are highly likely retained from the substratum due to the absence of appropriate superstratal replacements. Contrary to the common view, many basic terms associated with aspects of the Jewish religion are expressed by Slavicisms and not by loans from Hebrew and Judeo-Aramaic, the twin liturgical languages of the Jews. Conversely, Hebraisms are strikingly absent from the semantic fields of culture and religion.

5. A member of a language family might be relexification-based if it has much less inherited lexicon than other putative members. Yiddish has a reduced German component (and is especially poor in synonyms of German origin) compared to any single German dialect.

6. A language may be relexification-based if it lacks most of the native derivational morphology and allomorphic variants found in allegedly related languages. For exam-

ple, the German lexicon of Yiddish displays very few productive German derivational processes.

7. A language may be relexification-based if there is evidence of 'updating' or 'modernizing' the shape of older loans of one source *according to the sound changes in the donor language*. This fact suggests widespread bilingualism which often accompanies (cyclical) relexification. For example, Germanisms known to have been in Yiddish since the relexification periods almost always appear in contemporary Yiddish in Modern rather than Middle High German form.

8. A language may be relexification-based if morphological and semantic originality characterize almost exclusively only the lexifiers. The striking originality of the German component in Yiddish stands in sharp contrast to the relatively few innovative forms and meanings in the Slavic component of Yiddish. Also Hebroidisms outnumber original Slavisms as replacements for blocked Germanisms. These facts suggest that German and Hebrew were superstrata. Future studies should attempt to quantify these impressions.

9. A unique reflection of the Sorbian substratum is the parallel lexical gaps in Yiddish and Sorbian – where Yiddish used Hebraisms in place of Germanisms and Sorbian tended to borrow Germanisms. The parallelism between the distribution of Hebraisms in Yiddish and that of Germanisms in Sorbian finds no parallel even in Polabian, Czech and Polish, which were also very widely receptive to German enrichment, see e.g. Yiddish *recejex* 'murderer', *hargenen* 'to murder' < Hebrew ~ Upper Sorbian *mordar* 'murderer', *mordować* 'to murder' (since 1597) < German *Mörder, morden*, respectively.

Why would Slavic-speaking 'new Jews' have twice relexified? The most compelling explanation is that the confrontation of Germans and Slavs was resulting in the widespread erosion or extinction of Sorbian language, religion and culture, leaving the Slavic Jewish converts increasingly isolated from the bulk of the Germanizing Slavs. Relexification rather than a full shift to German might have allowed Slavic Jews symbolically to resist Christianization – a concomitant factor in the Germanization process, while still maintaining their unique linguistic and ethno-religious profile. Judaism would have been attractive to pagan Sorbs because no political commitments were involved (unlike Christianity which required espousal of German language, culture and political hegemony), and because Judaism offered Sorbs an opportunity to escape the status of slaves which was being broadly imposed on them by German and Viking slavers. The desire of the Turkic Khazar ruling class to preserve neutrality vis-à-vis the Byzantine Christians and the Baghdad Arab Caliphate, and the cultural influences of peripatetic Jewish merchants, initially made Judaism popular throughout Eastern Europe. It is widely assumed that the Ashkenazic Jewish immigrants from the Germano-Slavic lands to Eastern Europe vastly outnumbered the indigenous East Slavic-speaking Jews, but the facts of Yiddish suggest a Slavic-speaking majority

there. Recent genetic studies also offer some support for the linguistic conclusions (see Thomas et al. 2002; Behar et al. 2003 show that the Levites, a paternally inherited Jewish caste, display mutiple origins while the Cohanim, another, priestly, caste, display a common Near Eastern ancestry; the remaining 92% of the Ashkenazic Jews show affinities with Slavs).

3. Distinctive features of Yiddish relexification

The Ashkenazic Jews have undergone no less than six acts of relexification over the last eleven centuries: five acts of relexification of a Slavic language (a–e) and one act of relexification of Old (Semitic) Hebrew (f); → stands for 'became relexified to':

(a) Sorbian → Middle High German, Hebrew/Hebroid lexicon = Slavic Yiddish I (9th–12th centuries).
(b) Yiddish → Old Hebrew and Yiddish Hebroid lexicon = written Ashkenazic Hebrew (12th–19th centuries, if not earlier), a Slavic language with a predominantly Hebrew/Hebroid lexicon. (Written Ashkenazic Hebrew alaso enjoyed a spoken function as the basis of 'crypto-Yiddish' lexicons, possessing as much as 85% Hebraisms, used primarily in German-speaking lands.)
(c) Kiev-Polessian → Yiddish, Middle High German, Hebrew/Hebroid lexicon = Slavic Yiddish II (by the 15th century).
(d) Yiddish → Old Hebrew (as in [b]) but without many Yiddish Hebroidisms = spoken/written Modern Hebrew (late 19th century), a Slavic language with a predominantly Hebrew/Hebroid lexicon.
(e) Yiddish → Latinoid lexicon = Esperanto (late 19th century), a Slavic language with a Latinoid lexicon (the founder of Esperanto, Ludwik Zamenhof, was a native speaker of Yiddish).
(f) Old Semitic Hebrew → Middle High German lexicon = literal Yiddish Bible 'translations' (14th–18th centuries), unusual dialects of Old Semitic Hebrew with a predominantly Germanic lexicon.

Nine distinctive features of Yiddish relexification are exemplified by our dictionary entries:

1. Often, it is the morphologically least complex, or unmarked, member of the German paradigm that Yiddish selects, e.g. German *Berg* ∼ *Gebirge* = Upper Sorbian *hora* 'mountain', with no morphophonemic alternations; thus, Yiddish accepts only the simplex *barg* 'mountain'.

2. Yiddish pluralizes many German nouns with -(*e*)*n* in violation of German norms, but often in imitation of the Ukrainian/Belarusian 'pseudo-dual' number (where a noun following the numerals 2–3–4 takes the plural suffix but keeps the stress of the singular stem). This suggests that Yiddish once had a dual category.

3. The gender assignment of most German and Hebrew nouns in Yiddish dialects follows the gender of the Upper Sorbian/Kiev-Polessian translation equivalents rather than that of German and Hebrew nouns. Also, as the second component of compounds, nouns often change gender according to the gender of the Upper Sorbian or Kiev-Polessian translation equivalents; this is not typical of German compounds. Exceptions appear to be in the main demonstrably post-relexificational acquisitions.

If German has a single translation equivalent of several Slavic terms with various gender assignments, then relexification requires the selection of one of the underlying Slavic genders; a Yiddish noun rarely has the genders of all the Slavic equivalents. The choice of underlying Slavic gender may cast light on the inventory and meanings of the Slavic substratal lexicon at the time of relexification. For example, 'belly' in Ukrainian is now *žyvit, šlunok* m or *čerevo* n. The feminine gender of Lithuanian Yiddish *bojx* (~ masculine in other dialects) 'belly, abdomen' suggests the original substratal term prior to relexification was Kiev-Polessian (Ukrainian) *čerevo* n. Slavic feminine and masculine gender were originally assigned to the superstratal relexified German and/or Hebrew terms. Slavic neuter was automatically replaced by feminine gender for phonological reasons: Slavic neuter nouns often end in /o/, which becomes /e/ in Yiddish, where stress is rarely final. Unlike German, where final /e/ is not an unambiguous sign of gender, Yiddish assigns feminine gender to substratal Slavic neuters now pronounced with /e/, as well as to their relexified superstratal German and Hebrew replacements.

4. After the termination of the two relexification processes, Yiddish could borrow from all of its original component sources (as well as from new sources like Polish and Russian), without following substratal Slavic considerations; post-relexificational acquisitions can often be identified by formal/semantic clues. Positing relative chronologies for the acquisition of Hebrew elements in Yiddish is very often possible, since most Hebroidisms are by definition coinages motivated by relexification. For instance, Yiddish *mojre* 'fear; danger' continues Old Hebrew *mōrā'* 'fear', but the second meaning must be attributed to Upper Sorbian *strach* 'fear; danger' which was required to replace the blocked German *Gefahr* 'danger'. Synonymous Yiddish Hebraisms, such as *ejme* f and *paxed* m 'fear', do not also mean 'danger', probably because they were acquired either in the second relexification in the Kiev-Polessian lands, or in the post-relexificational period, and thus could not be affected by Upper Sorbian lexico-semantic parameters. For 'vomit', Old Hebrew provides *qī'* m (to replace the blocked Germanism), but Yiddish prefers the innovative *keje* f from the same root, possibly to match the feminine gender of Ukrainian *bljuvota*. It is also possible to posit cautiously the relexification stage during which a Germanism was most likely acquired, e.g. German *Anker* 'anchor' > Yiddish *anker* probably in the Sorbian lands, since Upper Sorbian *kótwica* would have presented no problem to relexifiers; however, in the Kiev-Polessian area German *Anker* might have been blocked by Ukrainian *jakir*, the Scandinavian cognate of German *Anker*. We can sometimes also postulate whether Germanisms accepted by Yiddish in the first relexification phase were dropped in the

2nd relexification phase in order to conform to the requirements of the Kiev-Polessian substratum. For example, Upper Sorbian *chować* 'to hide' should have licensed relexification to German *bergen* in the first relexification phase; yet, the latter term is missing (now) in Yiddish, perhaps because Kiev-Polessian (Ukrainian) offers cognate *berehty* 'protect, guard' (with *h* < **g*), thus providing grounds for blockage and elimination. Most Hebrew verbal material reaches Yiddish in the form of an indeclinable masculine singular participle, which must be combined with an inflected German auxiliary verb, see e.g. Yiddish *bojdek zajn* 'examine' (< Hebrew 'examining' + German 'be'). A Judeo-Belorusian example from c. 1600 with a Hebrew participle and the Slavic copula reads: *ja tebja estim mekadeš byl* (literally 'I you with-this [ring] sanctifying was') 'I sanctified you with this (ring)'. The identical conjugation is used in Islamic languages to integrate Arabic (and Iranian) elements. Hence, this grammatical rule may be from the Khazar (Turko-Iranian) or East Slavic substratum of Yiddish, i.e. it dates from the second relexification phase.

5. Yiddish frequently assists (a) in reconstructing aspects of German and Slavic linguistic prehistory, and (b) in positing the relative chronology of the diffusion of German words into Slavic. For example, Upper Sorbian *-mołwić* 'speak, talk', is now restricted to compounds with different, but related, meanings, e.g. *wotmołwić* 'answer', *wotmołwa* 'answer', *so zamołwić* 'forgive', *zamołwity* 'responsible', *so* (*wu*)*zamołwić* 'excuse (oneself); defend'. The Yiddish (and comparative Slavic) data suggest that these are relatively recent calques of German compounds formed from *Wort* 'word'. If Upper Sorbian *wotmołwa* 'answer', *wotmołwić* 'to answer' and *zamołwity* 'responsible' existed at the time of the first relexification phase, then there would have been no problem for Yiddish to relexify to German *Antwort*, *antworten* and *verantwortlich,* respectively – with identical component structure. Since Yiddish uses primarily Hebraisms, *čuve* or *entfer* 'answer' (the latter, significantly, < German with morpheme distortion, probably because Slavic languages do not express 'word' and 'answer' by identical or related roots), *axrajesdik* 'responsible', *xajev* 'guilty, responsible, liable', *memune* 'responsible person', Upper Sorbian probably was not yet following German practice in using a single root for these concepts at the time of the first relexification. Yiddish also does not follow contemporary Ukrainian practise in using 'know/tell', as in *vidpovidal'nyj* 'responsible', *vidpovid'* 'answer', *vidpovidaty* 'to answer' < *povidaty* 'tell' < *vidaty* 'know'. This fact suggests that the second Slavic substratum may have had limited impact on Yiddish.

6. Awareness of genetic relationships is shown by the fact that Yiddish relexifiers were able to transfer a meaning from a blocked German allomorph to another member of the allomorph set that was accepted in Yiddish, and have Hebraisms and retained Slavisms in the allomorph set. In the examples below, bold letters denote different forms of a common root. For example, German (**a**^{1}) *dicht* 'thick, firm, dense, compact', *Dichte* 'density' (16th c), *dichten* 'make tight, seal' (15th c) are unrelated to the set (**a**^{2}) *dick* 'thick, swollen, corpulent, stout', *Dicke* 'thickness, corpulence', *dicken* 'thicken', *Dickicht* 'thicket' (17th c) / (**b**^{2}) *gedeihen* 'prosper, thrive'. Yiddish accepts

(**a**^{1}) *dixt* 'plywood', *diklex* 'chubby', *gedixt* 'thick, dense, close, heavy', *gedixtkejt* 'consistency; density', *gedixteniš* 'thicket' (< German [**a**^{2}]) + (**a**^{2}) *dik* 'thick, fat, stout'/ (**b**^{2}) *gedajen* 'prosper, thrive'. Cognate (**a**^{2}–**b**^{2}) are not historically related to (**a**^{1}), despite similarity in sound and meaning, but the two paradigms appear to be partly joined in Yiddish due to Slavic considerations. Hence, German (**a**^{2}) exists in Yiddish, but not in the same distribution as in German. For German (**a**^{2}) *Dicke* 'thickness', Yiddish forms *greb* 'thickness' from *grob* 'thick, fat, coarse, rough, crude, rude, gruff, obscene' ~ German *grob* 'coarse, rough; rude; thick'. The late chronology of German *Dichte*, *dichten* and *Dickicht* suggests that the corresponding Yiddish terms may also be recent (for 'tighten', Yiddish uses other Germanisms). Whereas German uses (**a**^{2}) to form 'thicket', Yiddish selects (**a**^{1}). The changed distribution could be due to the underlying Upper Sorbian substratum, where one root covers both: (**a**^{1}) German *dicht* 'thick', *Dichte* 'density' = Upper Sorbian *husty; hustota*, *hustosć* and (**a**^{2}) German *dick* 'thick', etc., (*ver*)*dicken* 'thicken', *Dickicht* 'thicket' = Upper Sorbian *husty*; *zhusćić, hustnyć*; *husćina*, *husć*, respectively.

7. Often, when Upper Sorbian only overlaps semantically with part of a German root, the latter is used in Yiddish with only those meanings which are licensed by the Upper Sorbian term, e.g. the absence of a German root which combines the meanings of Upper Sorbian (*so*) *jednać* 'unite, unify; reconcile; negotiate'/*so dojednać* 'agree' obliges Yiddish to split, for the most part, the relexification between German and Hebrew roots, see Yiddish *farejnikn* (*zix*), *baheftn* (*zix*) 'unite, unify' (from German) alongside *maskem zajn, mašve zajn zix, mušve vern* (from Hebrew, but also *ajnštimen mit* from German) in the meaning of 'agree' (*so* and *zix* are reflexive particles).

8. Relexification can take more than one form simultaneously, i.e. speakers may differ on the appropriateness of Germanisms or in their preference for unrelexified Slavicisms or material from the secondary lexifier, Hebrew. Some of these choices account for contemporary Yiddish dialect differences.

9. Relexification to Hebrew differs radically from relexification to German, since, unlike German – the primary lexifier source – there was no attempt to accept a whole Hebrew paradigm. For example, German *Haß* 'hate' and *Häßlichkeit* 'ugliness' were both blocked in Yiddish, but the replacements are two separate Hebrew roots. This means relexification to Hebrew was piecemeal. It is rare to encounter Hebrew antonym sets in Yiddish, but not so German antonym sets.

4. Towards a blockage-based etymological dictionary

No Yiddish etymological dictionaries have ever been compiled. The dictionary which I am currently preparing will have a unique format that will allow the reader to appreciate immediately the unique nature of relexified languages and the details of the relexification process(es). The few etymological dictionaries and observations of relex-

ified languages that have appeared so far imitate traditional etymological formats. For example, the authors of etymological dictionaries of relexified Caribbean creole languages see as their primary task the collection of unique substratal relics and regional or archaic superstratal elements. This approach creates the erroneous impression that the distinctive feature of Caribbean creoles is lexicon not current in the European lexifiers. Unfortunately, the English creole dictionaries of F.G. Cassidy and R.B. LePage (1979) and R. Allsopp (1996) both ignore the bulk of the European lexical corpus shared by the latter and Caribbean creoles, which they seem to regard as 'non-native' and irrelevant, as well as the possibility of predicting and motivating superstratal blockage altogether (see Section 2 for a description of blockage), which is the very hallmark of relexified languages.

Creole and Yiddish lexicographers assume that there is little point in etymologizing the vocabulary shared by each language with its lexifier language. I agree that there is little that one could say about Yiddish *frage, freg,* (archaic) *frog* 'question' that has not already been stated about German *Frage* in a German etymological or dialect dictionary. However, the job of an etymological dictionary of a relexified language should be to explain *why* Yiddish borrowed Middle High German *vrëge* (see also German dialectal *frägen*) and Middle High German *vrâg*(*e*) 'question' (Modern German *Frage*), and why the latter have to coexist with synonymous *šajle, kaše* < Hebrew, i.e. to *show how the substratal language determines the selection and calibration of superstratal lexicon.*

The differences between an etymological dictionary of a relexified and non-relexified language are major. An etymological dictionary of a non-relexified language typically (a) reconstructs a proto-form; (b) lists cognates; (c) treats only and all the extant forms of the language in question in a single synchronic crosscut; (d) arranges entries alphabetically by root (for a typology, see Malkiel 1975). The traditional dictionaries and etymological discussions of Yiddish derive Slavisms (c. 10% of the lexicon) indiscriminately from the closest-sounding source, reconstruct proto-Yiddish Germanisms (c. 75%), and treat Hebrew/Hebroid (c. 15%) as a single Semitic lexifier source.

An etymological dictionary of a relexified language, as proposed here, has (a) no reconstructed proto-forms (the user is referred to etymological dictionaries of the sub- and superstratal languages, in our case German and Hebrew; Upper Sorbian, Ukrainian, Belarusian); (b) does not need to list cognates of the substratal and superstratal words, since these data are available in etymological dictionaries of the respective languages; and (c) an etymological dictionary of a relexified language has to provide information not normally expected in the historical-etymological dictionary of a non-relexified language. It has to examine the *entire* corpus of the German relexifier dialect(s) postulated at the time of the relexifications, i.e. both blocked and accepted Germanisms and not just the superstratum German corpus actually found in Yiddish; and it has to try to explain why certain German vocabulary was blocked and to posit the likely underlying pre-relexified West and East Slavic terms that determined the selection and calibration of German and Hebrew phonetic strings during the two relexification phases. *The citation of 'non-existing' vocabulary should be a distinctive re-*

quirement of an etymological dictionary of any relexified language. (d) An etymological dictionary of a relexified language arranges entries by sets of genetically or semantically related roots (synonyms, antonyms), rather than by individual roots. (e) The Yiddish etymological dictionary will identify the three major components of the Yiddish lexicon in terms of pre-relexificational, relexificational and post-relexificational strata. By suggesting how to distinguish retained Slavic substrata from post-relexificational superstrata, and by explicitly distinguishing (for the first time) genuine Semitic Hebraisms from Yiddish 'Hebroidisms', our dictionary can constitute a major contribution to Hebrew (especially relexified 'Slavic Modern Hebrew') lexicography. Hopefully, the idea of taking *the primary superstratal lexifier (German) as the point of departure* in this dictionary of a relexified non-creole language will serve as a useful model for creole lexicography. The superstratum is the point of departure since it constitutes the lexical outcome of relexification; it goes without saying that the entries will also note the etymological and derivational links characterizing the substratal Slavic lexicon.

Post-relexification loans from Hebrew, Slavic and German should present no special problem for the historical lexicographer of Yiddish, though no dictionary of Yiddish distinguishes between relexification and straightforward loans from these three sources. Unfortunately, all Modern Hebrew dictionaries also subsume native and non-native Hebrew elements in a common listing. The on-going publication of Medieval Latin dictionaries according to the native language of the scribe is a model that Hebrew lexicographers ignore at their own peril. Unspoken Medieval Polish, Hungarian, English Latin, etc. are relexifications of Polish, Hungarian, English, etc. to Latin lexicon and thus are genetically distinct from native spoken and written Latin. What is needed are separate dictionaries of (a) 'native Semitic Hebrew' (obsolescent for centuries before becoming finally obsolete by the 3rd century A.D.), (b) 'non-native non-Semitic Hebrew' (e.g. 'Modern Israeli Hebrew', which is Yiddish, *qua* Slavic, relexified to real Semitic Hebrew lexicon: see Wexler 1990) and (c) 'non-native Semitic Hebrew' (e.g. the Hebrew written by native speakers of Arabic or Neo-Aramaic).

There are two advantages to my proposed format: (1) In already verified cases of relexification, *the superstratal lexicon licensed for acquisition is largely predictable*. Since much German vocabulary is blocked by the semantic and syntactic parameters of the substratal Slavic grammars, the German component of Yiddish is, not surprisingly, considerably smaller than that of any German lexifier. The lexical gaps caused by the blockage of large amounts of German data have to be filled by Germanisms that are acceptable, Hebraisms (a secondary lexifier) or, in partially specifiable semantic domains, by unrelexified Slavisms; where Hebraisms are lacking, Yiddish invents large numbers of formal and/or semantic Hebroidisms (only relexification can explain why Yiddish Hebraisms greatly exceed the 'Hebrew' corpus of all other Jewish languages). The format adopted here (shown in the two sample examples below) explicitly displays the process of relexification word by word (a) by showing the extent and location of blockage, and (b) by motivating the use of replacements. (2) In languages where relexification is so far only suspected, *my format provides an ideal diagnostic test for relexification*.

5. Sample entries

In the two sample entries from the Yiddish etymological dictionary now in progress that are presented below, the order of each superordinate entry is (i) Yiddish allomorphs of a common German root + (ii) Germanisms not recommended for Yiddish + (iii) Hebroidisms (predominantly generated during relexification) and Hebraisms (acquired both via relexification and as post-relexificational loans) + (iv) Germanoidisms and Slavisms (un-relexified retentions and post-relexificational loans + (v) other Yiddish German(oid)isms cited in the entries. Following the German paradigms, after the symbol //, are all other Germanisms cited in the entries.

Abbreviations used in the sample entries below are: ag = agentive, Ar = Judeo-Aramaic, Br = Belarusian, (C)Sl = (Common) Slavic, f = feminine, G = German, He = Hebrew, KP = Kiev-Polessian, m = masculine, MH = Middle High, n = neuter, plt = plurale tantum, USo = Upper Sorbian, Uk = Ukrainian, Y = Yiddish. Symbols placed before examples are: ▲ = Hebraism; ▲ = Hebroidism; ◙ = Slavism; ◙ = Slavoidism (thus enabling the reader to appreciate the extent of acquisition/blockage of Germanisms); absence of a symbol denotes a Germanism. ¶ = a post-Middle High German term or post-relexificational element in Yiddish (i.e. after c.1500); ♣ = Germanoidism in Yiddish unattested in German modeled on Slavic; º = not recommended for standard Yiddish; © = cognate, ® = 'relexification, relexify', ®'ied = 'relexified'. ®1 = the first relexification phase from Sorbian (9th–12th centuries); ®2 = the second phase from East Slavic (11th century[?]–1500s at the latest). Cognate forms are separated by the symbol /; non-cognate forms by the symbol +. German terms created after the 16th century are not cited unless they entered Yiddish post-relexificationally. The notation '(< [**a**])' means that the allomorph in use comes from another part of the paradigm – a fact which reflects the sensitivity of relexifiers to the constraints of blockage and their awareness of etymological relationships. In the actual dictionary entries, all Germanisms cited in the text will also be listed; for ease of use, all Yiddish examples are listed alphabetically and examples from all other languages are listed in indices. Hebrew and Judeo-Aramaic etyma are given as transliterated roots, which allows us to avoid imposing an anachronistic pronunciation norm.

(*German*) **ajnkojf(er), ajnkojfn, farhandlen, farhandlungen, farkojf(er), farkojflex, (far)kojfn, handl(en), komerc(iel), ojfkojfn, ojsfarkojf, ophandlung** + (*blocked German*) **º(unter)handlung** + (*Hebrew/Hebroid*) **▲farmisxert, ▲farmisxerung, ▲hobn/maxn maxzokes mit, ▲kinjen, ▲knie, ▲kojne, ▲kojnete, ▲magemase, ▲maxzokes, ▲mecie, ▲mexire, ▲misxer, ▲saxren, ▲sojxer, ▲sxojre, ▲xiber** + (*Germanoid*) **♣far-, ♣farkojfn zix far geld, ♣ojskojfn (zix fun), ♣onkojfn, ♣opkojfn, ♣unter(kojfn)** + (*Slavic/Slavoid*) **◙farkojferke** [+ (*other German/Germanoid*) **arbet, falš, farker, farmegn, krum, šlext umgang**]

G: (a) ¶**aufkaufen** 'buy up', ¶**Ausverkauf** 'clearance', **Einkauf** 'purchase', **Kauf** m 'buying, purchase; bargain', **kaufen** 'buy', **Kaufmann** m 'merchant', ¶**Kaufmannsgut** n 'ware(s)', ¶**Verkauf** m 'sale' (16c), **verkaufen** 'sell'/ **(b) Käufer(in)** m (f) 'buyer',

¶**Verkäufer(in)** m (f) 'seller' (17c) (< MHG [**a**] **verkoufære** m), G (**ver**)**käuflich** ¶'for sale, marketable' (< MHG [**a**] **kouflich** 'profitable, disposable') + (**c**) G **Handel** m 'trade, commerce', **handeln** 'to trade', **abhandeln** 'beat down (price); treat, discuss', ¶**Abhandlung** f 'treatise, essay; dissertation', ¶**behandeln** 'to treat, manage, deal with; handle', ¶**Behandlung** f 'treatment, management; use, handling' (16c), **misshandeln** 'mistreat, manhandle', ¶**Misshandlung** f 'mistreatment', ¶**unterhandeln** 'negotiate', ¶**Unterhandlung** f 'transaction' (16c), **verhandeln** 'discuss, treat; sell, negotiate; try (in court)', **Verhandlung** f 'discussion, treaty; negotiation; trial'/ (**d**) **Händler** m 'dealer, trader' + (**e**) MHG **sellen** 'to sell' + (**f**) G ¶**Kommerz** m, n 'commerce' (16c), ¶**kommerziell** 'commercial' (19c)// *Arm, bestechen,* MHG *mieten, salben,* G *schachern,* MHG *smirwen, stëchen, überkomen* >
Y: (**a**) **kojfn** 'to buy', **ajnkojf** 'purchase', **ajnkojfer** m '(professional) buyer', **ajnkojfn** 'to buy; shop (for)', ¶**farkojf** m 'sale(s)', ¶**farkojfer** m 'seller' (< G [**b**]), **farkojflex** 'negotiable; venal', **farkojfn** 'to sell', ¶**ojfkojfn** 'buy up', ¶**ojsfarkojf** m '(clearance) sale' + (**c**) **handl** m 'trade, commerce; dealings', **handlen** 'to buy', º**handlung** 'deed, action', **farhandlungen** plt 'negotiations, proceedings', **farhandlen** 'sell; negotiate', ¶**ophandlung** 'treatise, discourse, dissertation', ¶º**unterhandlung** f + (**f**) ¶**komerc** m 'commerce', ¶**komerciel** 'commercial'.

There are six reasons why I suspect that originally Y did not accept either G *kaufen* 'to buy' or *verkaufen* 'to sell':

(i) The G'ism (ultimately < Latin) was probably borrowed by CSl, and USo *kupić* 'to buy' is formally close (thus, the use of the He'ism for parts of the G paradigm – see below).

(ii) The Sl languages lack the vocalic alternation that characterizes G (**a–b**), see e.g. (**a**) USo *kupić* 'to buy'/ (**b**) *kupc* ~ *kupowar* m 'buyer'.

(iii) Semantic disparallels between G and Sl may also have contributed to initial blockage, see e.g. MHG (**a**) *koufen* ~ *keufen* could mean both 'buy' and 'sell' (while MHG *verkoufen* ~ *verkeufen* meant variously 'sell; surrender, sacrifice; fritter away'). Semantic disparallels between G and Sl also obtained in the KP area, see e.g. Uk *kupyty bidu sobi* literally 'buy trouble to-self', i.e. 'bring trouble on oneself'. Failure to acquire G *Kaufmann* 'merchant' could be explained by the fact that CSl **kuьcь* m was ambiguous – 'merchant; buyer'.

(iv) USo derives *kupc* m 'buyer' < *kupić* 'to buy'; the former has given rise to a new verb *kupčić* 'haggle; speculate'. The inability to ® the pair USo *kupić*/*kupčić* (assuming it already existed in ®1) to a single root in G may be the reason why Y created the colloquial non-periphrastic ▲**saxren** 'do business' < He *sḥr* m 'trade' – perhaps originally in the now unattested meaning *'haggle' (this meaning is attested with G *schachern*, known since 1726).

(v) Only blockage would account for the wealth of Y He'isms: for (**a**) 'bargain' and 'purchase', see also Y ▲**mecie**, ▲**knie** f, respectively (< He *mcj'h, qnjh*); for (**a**) 'merchant', see Y ▲**sojxer** m, along with the related ▲**sxojre** f 'ware(s), goods', ▲**misxer** m 'trade, commerce, business', ▲**farmisxert** 'commercialized',

▲**farmisxerung** f 'commercialization' (< He *swħr, sħwrh, msħr*); the use of ♣**far-** ~ Uk *za-* in Uk *zatorhuvaty* 'begin to earn a living from trading' < *torhuvaty* 'to trade, sell'. For (**b**), Y uses ▲**kojne** m, ▲**kojnete** f 'buyer' (related to ▲**knie** < He *qwnh* + Ar *-t'* f ag). From the same He root Y has also acquired ▲**kinjen** m 'acquisition; (spiritual) property' (< He *qnjn*), coexisting with **farmegn** n 'possession, estate'.

(vi) Another sign of possible initial blockage is that Y dialectally has another pair of terms, **handlen** 'to buy'/**farhandlen** 'to sell'.

While the above reasons suggest a late acquisition of the G'isms, either in ®1 or thereafter, there is still one fact which could theoretically have licensed full ® in the G-So lands: MHG expressed the opposition of 'buy' and 'sell' in two ways: *koufen* ~ *keufen* 'to buy'/*verkoufen* ~ *verkeufen* or *sellen* 'sell'; the latter pair could have licensed early ®, given USo *kupić*/*předać,* Uk *kupyty*/*prodavaty*. This may not be a weighty argument, due to the semantic heterogeneity of both MHG *koufen* ~ *keufen* and *sellen* (also 'hand over, deliver; sell retail'; see point iii above).

Y ♣**unterkojfn** 'to bribe' with ♣**unter** 'under' ~ Uk *pidkupyty* < *kupyty* 'to buy' + *pid-* 'under'; G *bestechen* ¶'to bribe' was too new to reach Y and MHG *mieten, salben, smirwen, stëchen, überkomen* 'to bribe' were all blocked in Y. Y (**a**) ¶**farkojfer** m 'seller' and ¶◙ **farkojferke** f 'seller' (the latter with an innovative use of Sl *-ka* f ag vs. Uk *prodavnycja, prodavčycja* f < *prodaval'nyk, prodavec'* m) are recent loans < G; Y ♣**farkojfn zix far geld** 'prostitute oneself' (literally 'sell oneself for money') ~ Uk *prodavatysja* 'sell, be sold'; ♣**onkojfn** 'heap; stack; amass' ~ Uk *nakupyty* 'buy in large quantities, make many purchases', ♣**ojskojfn** 'buy up; ransom, redeem', ♣**ojskojfn zix (fun)** 'atone (for)' ~ Uk *vykupyty* 'redeem, random, buy back, repurchase', *vykupytysja* 'be ransomed, redeem oneself', ♣**opkojfn** 'buy (from)' ~ Uk *vidkupyty* 'redeem, ransom, repurchase', *vidkupytysja* 'redeem oneself'. Y ▲**mexire** f 'sale, act of selling' (< He *mxjrh*) is required because Uk distinguishes between *prodaž* 'selling of goods' and *zbut* m 'transfer of goods'.

Y (**c**) **handl** in the meaning 'dealings' coexists with ▲**maxzokes** plt 'dealings; temporizing', ▲**hobn/maxn maxzojkes mit** 'bother, fuss with' (< He *mħzwqt*), ▲**magemase** m 'dealings; (social) intecourse' (< He *mg'-wmś'*); in the latter meanings, see also Y **farker** m, °**umgang** m. Y ▲**xiber** m 'treatise; addition (mathematics), composition' (< He *ħjbwr*) coexists with **arbet** f 'written composition'. For G *miss-* 'mis-' in (**c**), Y uses an adverb, e.g. **šlext** 'badly', **falš** 'falsely', **krum** 'poorly, wrong'.

(*Hebrew/Hebroid*) ▲**vald-xazer,** ▲**xazeraj,** ▲**xazer(-flejš),** ▲**xazer-štal,** ▲**xazere,** ▲**xazerte** + (*Germanoid*) ♣**(runclen) zix** + (*Slavic/Slavoid*) ◙**brozde,** ◙**gare,** ◙**kaban,** ◙**ljoxe** [+ (*other German/Germanoid*) **vald**]
G: (**a**) **Sau** f 'sow'/ (**b**) **säuisch** 'piggish; obscene'/ (**c**) **Schwein** 'pig', **Schweinefleisch** n 'pork', Schweinestall f 'pigsty' + (**d**) **Eber** m 'boar' + (**e**) **Ferkel** n 'piglet'/ (**f**) **Furche** f 'furrow', **furchen** 'to furrow', **furchig**, **zerfurcht** 'wrinkled, furrowed' + (**g**) **Frischling** m 'young wild boar' + (**h**) MHG **lôse** f 'breeding sow' + (**i**) **tocke** f 'sow' + (**j**) G **Egge** f 'harrow', **eggen** 'to harrow'// (*die Stirn) runzeln*.

The entire German corpus is blocked in Yiddish. For G (**a-e, g-i**), Y uses ▲**xazer** m with innovative ▲**xazerte** (< He *ħzjr* + Ar *-t'* f ag) and ▲**xazere** f (with a He f suffix), or ◙**ljoxe** 'sow' < Uk *l'oxa* f (< MHG [**h**] with characteristic expressive change of /s > x/), ▲**xazeraj** n, f 'filth, mess, dirty trick', ▲**vald-xazer** (with **vald** 'forest'), ◙**kaban** 'wild boar' (< Uk [*dykyj*] *kaban* m 'castrated pig' [literally 'wild pig']). Y ▲**xazer(-flejš)** n 'pork' has optional use of G 'meat', following compound G *Schweinefleisch* rather than simplex Uk *svynyna* f 'pork'. Y ▲**xazer-štal** f, ♣m replaces the blocked G (**c**), with the optional m gender < Uk *xliv, svynarnyk, svynynec', svynjušnyk* m; however, a compound noun in Y matches G rather than Sl usage. A He'ism would be appropriate for G (**a-c**) because the G and Sl terms are ©s, see (**a**) USo *swinica,* Uk *svynomatka* f 'sow'/ (**b**) USo *swinski,* Uk *svynjačyj*/ (**c**) USo *swinjo* n, Uk *svynja* f. Significantly, Y retains the Sl practice of using a common root without the morphophonemic alternation and lexical variety of G. G (**d**) is blocked because of © Uk *vepr* m; G (**e-f, j**) are lacking because the Sl lexicon requires multiple roots, see (**e**) USo *proso* n, while denoting 'furrow' and 'harrow' by related and formally similar terms: (**f**) USo *brózda*, Uk *borozda* 'furrow'; the latter has a variant with *-n-* in Uk *borozna,* Br *barazna* f/ (**j**) USo *bróna*, Uk *borona* 'harrow'. For G (**f**) '(to) furrow', see Y ◙**brozde** 'furrow', **gare** ◙ 'furrow'; 'groove' < Uk *borozda* ~ *borozna* 'furrow', Uk *gara* 'groove; socket'; for G (**j**), see Y ◙**brone** f (< Uk with deletion of the initial unstressed syllable in Y [**f, j**] or < USo). G (**f**) *furchig, zerfurcht* 'wrinkled, furrowed' are unacceptable to Y, since no single Sl root can denote both 'furrow' and 'wrinkled', but see Y ♣**runclen zix** < G *runzeln* 'to wrinkle, pucker', *die Stirn runzeln* 'frown, knit one's brow' (Y reflexive ♣**zix** ~ Uk reflexive *-sja* in *morščyty*[*sja*]).

The proof of relexification in the genesis of Yiddish must ultimately come from two kinds of research (as outlined in the first two diagnostic tests): (a) The grammar and phonology must be shown to be predominantly Slavic in origin, in opposition to the bulk of the lexicon; and (b) all of the lexical components of Yiddish must be 'predicted' and the blockage by Yiddish speakers of large numbers of available Germanisms must be convincingly motivated. The on-going compilation of the first etymological dictionary of Yiddish (and for that matter of any relexified language) will provide full support for the second research requirement.

References

Allsopp, R. (1996). *Dictionary of Caribbean English Usage*. Oxford: OUP.

Behar, D. M., Thomas, M. G., Skorecki, K., Hammer, M. F., Bulygina, E., Rosengarten, D., Jones, A. L., Held, K., Moses, V., Goldstein, D. B., Bradman, N., & Weale, M. E. (2003). Multiple origins of Ashkenazic Levites: Y chromosome evidence for both Near Eastern and European ancestries. *American Journal of Human Genetics, 73*(4), 768–779.

Cassidy, F. G. & LePage, R. B. (1979). *Dictionary of Jamaican Creole English* (2nd edition). Cambridge: CUP.

Geller, E. (2001). *Warschauer Jiddisch*. Tübingen: Niemeyer.

Horvath, J. & Wexler, P. (1994). Unspoken 'languages' and the issue of genetic classification: The case of Hebrew. *Linguistics, 32*(2), 241–269.

Horvath, J. & Wexler, P. (1997). Relexification: Prolegomena to a research program. In J. Horvath & P. Wexler (Eds.), *Relexification in Creole and Non-creole Languages.* (*With special reference to Modern Hebrew, Haitian Creole, Romani and Rumanian.*) (pp. 11–71). Wiesbaden: Harrassowitz.

Katz, D. (1985). Hebrew, Aramaic and the rise of Yiddish. In J. A.Fishman (Ed.), *Readings in the Sociology of Jewish Languages* (pp. 85–103). Leiden: E.J. Brill.

Lefebvre, C. (1986). Relexification in creole genesis revisited: The case of Haitian Creole. In P. Muysken & N. S. H. Smith (Eds.), *Substrata Versus Universals in Creole Genesis* (pp. 279–300). Amsterdam: John Benjamins.

Lötzsch, R. (1990). *Jiddisches Wörterbuch.* Leipzig: Bibliographisches Institut.

Lötzsch, R. (1996). Interferenzbedingte grammatische Konvergenzen und Divergenzen zwischen Sorbisch und Jiddisch. *Sprachtypologie und Universalienforschungen, 49*(1), 50–59.

Malkiel, Y. (1975). Etymology and modern linguistics. *Lingua, 36*, 101–120.

Manaster Ramer, A. (with the assistance of M. Wolf). (1996). Yiddish origins: The Austro-Bavarian problem. *Folia Linguistica Historica, 17*(1–2), 193–209.

Thomas, M. G., Weale, M. E., Jones, A. L., Richards, M., Smith, A., Redhead, N., Torroni, A., Scozzari, R., Gratrix, F., Tarekegn, A., Wilson, J. F., Capelli, C., Bradman, N., & Goldstein, D. B. (2002). Founding mothers of Jewish communities: Geographically separated Jewish groups were independently founded by very few female ancestors. *American Journal of Human Genetics, 70*, 1411–1420.

Vajnrajx, M. (1973). *Gešixte fun der jidišer šprax* 1–4. New York NY: YIVO. Partially translated as *The History of the Yiddish Language*, 1980. Chicago IL: Chicago University.

Weinreich, U. (1958). Yiddish and Colonial German in Eastern Europe: The differential impact of Slavic. *American Contributions to the International Congress of Slavicists, 4*, 369–421. The Hague: Mouton.

Wexler, P. (1983). Is Karaite a Jewish language? *Mediterranean Language Review, 1*, 27–54.

Wexler, P. (1990). *The Schizoid Nature of Modern Hebrew. A Slavic language in search of a Semitic past.* Wiesbaden: Harrassowitz.

Wexler, P. (1991). Yiddish – the fifteenth Slavic language. *International Journal of the Sociology of Language, 91*, 1–150, 215–225. (Accompanied by critiques by Bernard Comrie, James R. Dow and Thomas Stolz, Paul Glasser, Neil G. Jacobs, David F. Marshall, Gunter Schaarschmidt, Heinz Schuster-Šewc and Edward Stankiewicz.)

Wexler, P. (1992). *The Balkan Substratum of Yiddish. A reassessment of the unique Romance and Greek components.* Wiesbaden: Harrassowitz.

Wexler, P. (2001). De la relexification en yiddish: Les Juifs, les Sorabes, les Khazars et la région de Kyjiv-Polissja. *Canadian Slavonic Papers/Revue Canadienne des Slavistes, 43*(4), 533–556.

Wexler, P. (2002a). A German-Slavic bilingual unknowingly knows Yiddish and Modern Hebrew quite well: Why and what? *Welt der Slaven, 47*(1), 57–84.

Wexler, P. (2002b). *Two-tiered relexification in Yiddish.* (*The Jews, Sorbs, Khazars and the Kiev-Polessian dialect.*) Berlin: Mouton de Gruyter.

Wolf, S. A. (1956). *Wörterbuch des Rotwelschen. Deutsche Gaunersprache.* Mannheim.

PART II

Variation

A fresh look at habitual *be* in AAVE

Chris Collins
New York University, USA

In this paper, I describe the use of agentive *be* ('If you don't be careful, you will be caught') in informal American English. I will show that agentive *be* has largely the same syntactic behavior as habitual *be* in AAVE (African American Vernacular English). Based on these similarities, I will conclude the paper by raising a number of questions about the origin of habitual *be* in AAVE.

1. Introduction[1]

In this paper, I will describe the use of agentive *be* in informal American English, basing my analysis primarily on my own intuitions (I-language) and sentences that I have found on the Internet. I will show that agentive *be* has largely the same syntactic behavior as habitual *be* in AAVE. Based on these similarities, I will conclude the paper by raising a number of questions about the origin of habitual *be* in AAVE.

Consider the contrast between (1b) and (2b):

(1) a. If you are not careful, you will be caught.
b. If you don't be careful, you will be caught.

(2) a. If you are not seen, you will escape.
b. *If you don't be seen, you will escape.

The fact that (2b) is ungrammatical while (1b) is not seems related to the fact that one can be deliberately careful, but it is less likely that one can be deliberately seen (especially in the context of an escape). In other words, there seems to be a particular verb *be* whose semantics implies deliberateness. Henceforth, I will call this agentive *be* in order to distinguish it from other uses of *be*.

1. I would like to thank Arthur Bell, Richard Kayne, Wayne Harbert, Eric Potsdam and John Whitman for comments on earlier drafts of this paper. By e-mail correspondence, Lisa Green gave me valuable advice on certain aspects of habitual *be*. I would also like to thank two reviewers who suggested some important references on AAVE, and whose insightful comments have helped me to improve the paper.

The acceptability of (1b) is surprising in light of the fact that *be* does not normally permit do-support. The reason for this is that *be* normally raises to Infl, if Infl is not occupied by a modal auxiliary. The classical paradigm that illustrates this point is given below (see for example, Pollock 1989: 398):

(3) a. John isn't happy.
b. *John doesn't be happy.
c. John shouldn't be happy.

Sentence (3a) illustrates that the auxiliary verb *be* moves to Infl, raising over negation. The fact that *be* can move to Infl accounts for the unacceptability of (3b) under the assumption that do-support is a Last Resort mechanism, used only when necessary. Lastly, if Infl is filled by the modal auxiliary *should* as in (3c), *be* cannot undergo movement to Infl (since Infl is already filled).

In section two of this paper, I will present a range of examples of agentive *be* in different contexts. In section three, I will show that agentive *be* cannot have any overt inflection. In section four, I will distinguish agentive *be* from the imperative and subjunctive. In section five, I will discuss how agentive *be* fits into a theory of verb movement. In section six, I will discuss the syntactic similarities and differences between habitual *be* in AAVE and agentive *be*. In section seven, I outline one way in which agentive *be* could have influenced the development of habitual *be* in AAVE.

2. The distribution of agentive *be*

Since judgments involving agentive *be* are insecure for some people, and since the construction is definitely informal, I have performed internet searches (from May 2000 to January 2005) in order to find examples and to present a fuller picture. I used various search engines, including Google and Fast. In doing the internet searches, the basic methodology is to search for a phrase such as 'if you be careful' (and variants such as 'if you don't be careful', 'if he doesn't be careful', etc.). The search engine then produces a list of web pages where this phrase is found. Next I analyzed each page to find the first occurrence of the phrase (such as 'if you be careful') on that page. For reasons discussed below, I exclude all pages that deal with archaic English (e.g., early modern English, biblical English, or even the English from the writings of Charles Dickens) or AAVE (African-American Vernacular English). The sentences I found on the internet are all acceptable or nearly acceptable in my personal dialect of informal English. In doing these searches, I noticed that the number of examples found for any given search (e.g., 'if you be careful') goes up from one time to the next. Furthermore, URLs become inactive frequently, perhaps in part due to the informal nature of the internet sites where the examples are found. Therefore, I have not included any URLs for the examples given in the paper. Most of the URLs are still active, and to find them, the reader just needs to enter the example into Google.

I give the examples that I found on the internet in (4) through (30) (except (20), which was provided to me by Wayne Harbert). I have retained the punctuation found on the web pages (including lower case letters, where upper case letters are expected and vice versa) to give a flavor for how informal the construction is. The first group of examples illustrates the range of adjectival predicates found with agentive *be*. I do not claim that this group of adjectives is an exhaustive list. These examples also illustrate that agentive *be* can be found with or without negation. When there is negation, there is do-support (see (5)):

(4) If you be careful, it's neat to ride in the rain. Old lady thought I was nuts!!!

(5) If you don't be careful and vigilant, that person they talk to one day could be you.

(6) The salamanders will ignore you if you be quiet and just watch.

(7) If you be nice to people they'll be nice back.

(8) BUT, if you be honest, and be strong-people will respect you for your decisions.

(9) But if you be like that, you won't get the chance to see one of the best comics that I've ever seen!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

(10) You will get nowhere if you be mean.

(11) Perhaps if you be more specific in describing this problem...

The second group of examples shows how agentive *be* appears with first person and third person subjects (as well as the second person subjects seen in the above group of examples):

(12) If I don't be careful, I am going to know the words to YPIMP by heart.

(13) If we don't be careful, we're gonna be the angriest and then we won't be happy anymore.

(14) I keep telling him that he is going to be in a wheelchair if he doesn't be careful.

Agentive *be* is also possible in when-clauses:

(15) Be nice day is when you be nice to a lot of people.

(16) When you be careful enough to check during the install procedure. . .

(17) I find it usually only happens when we be nice and activate the phone in the store.

It is possible (perhaps marginally) to find agentive *be* in matrix declarative clauses:

(18) We be nice, we be nice, we be nice; then we kill them.

(19) Used to be that i thought i only be nice to those who deserve it.

(20) We be nice when we're trying to impress the teacher. (Wayne Harbert, personal communication).

Agentive *be* can also be found in the past tense (with the auxiliary verb *did*):

(21) If he didn't be careful, he would almost feel like he was falling in love with her.

(22) I showed them some fish that could hurt them if they didn't be careful.

(23) 'And did you be nice?' Santa asked sharply.[2]

Lastly, agentive *be* can be used with DP predicates:

(24) Did you be yourself, or did you change your voice to fit the characters?

(25) Successful leaders exude confidence, which can only happen when you be yourself.

(26) If I be myself, people get upset, cry, get pissed off or leave me.

(27) Please? Do it again, I just want to figure out how you did that, meaning how did you *be* my friend.

(28) If you be a good boy and go back to Havana, we'll see what kind of aid you get.

(29) How do you be a model for your kids?

(30) How do you be a good Catholic parent when society pushes away religion?

The adjectives that can be used with agentive *be* are the type that can be interpreted eventively and agentively, as shown by their use in the progressive, which forces an agentive interpretation (see Partee 1977: 32):

(31) a. I am being careful, vigilant, nice, quiet, honest, strong...
b. Why are you being like that, mean, so specific...

In fact, it is reasonable to assume that the agentive *be* found in the examples in (4–30) is the same as the active *be* discussed by Partee in connection with the progressive.[3] The predicate nominals that are used with agentive *be* above can also be found in the progressive:

(32) a. I am just being myself.
b. I am just being your friend.
c. I am being a good boy.

---

2. See Becker (2004: 404) for a similar sentence drawn from her own dialect. She notes that the question 'Did you be good today?' is less acceptable than 'Were you good today?' but more acceptable than 'Did you be smart today?' I agree with these judgments.

3. Green (2000: 15) analyses a sentence such as 'Richard is being nice' as having 'event argument associated with stage-level predicates'. Additionally, on my analysis, the subject is interpreted as the agent (and not, for example, the theme) of the event argument. In other words, 'Richard is being nice' cannot mean that Richard is becoming nice, which would also be an eventive interpretation.

d. I am being a model for my kids.
e. I am being a good Catholic parent.

Subjects and predicate adjectives that are not consistent with an agentive interpretation cannot be used with agentive *be*.

(33) a. If you be noisy, you will bother your sister.
b. *If the party be noisy, it will bother your sister.
c. You are being noisy. (agentive, active interpretation only)
d. *The party is being noisy.

(34) a. If you be careful, you will not fall.
b. *If you be tall, you will be able to reach the top shelf.
c. You are being careful. (agentive, active interpretation only)
d. *You are being tall.

3. The overt inflection constraint

The following example with a third person singular subject is unacceptable to me.

(35) a. *If he be careful,...
b. *If he be's/bes/bees careful,...

A search of the internet yielded results consistent with the grammaticality judgments in (35). A search of 'If he doesn't be careful' turned up the following example among others (for example, (14) above):

(36) If he fights, it's out of desperation. If he doesn't be careful, he's going to get hurt in the ring.

On the other hand, a search of 'if he bes careful' (spelled b-e-s) or of 'if he bees careful' (spelled b-e-e-s) or of 'if he be's careful' (spelled b-e'-s) turned up no results at all. A search of 'if he be careful' turned up around 10 results, but all of them were in archaic language (e.g., *Dickens's Dictionary of London*, by Charles Dickens, Jr.), and so represent the formal subjunctive (see Section 4 below).

The sentence in (35a) is unacceptable, because the subject is third person singular, and the third person singular agreement morphology has not been spelled-out (compare 'John runs' versus *'John run').[4] (35b) seems to be ruled out because the verb

4. I assume that the lack of agreement with modal verbs ('John would/*woulds run') is related to the fact that they are in Infl (e.g., modals never follow negation). Since Modal-s would not involve affix hopping, the lack of Modal-s suggests that it is only necessary to spell-out [+3sg, +Pres] as *-s* when there is affix hopping (see also Kayne 2000: 198 who gives a more general condition).

be inflected for third person singular[5] must be spelled out as *is* not *bees*. The relevant constraints are the following:

(37) a. [$_{\text{Infl}}$ be [+3sg, +Pres]] → [Iz] (morphological spell-out)
b. [be [3sg, present]] → [Iz] (morphological spell-out)
c. The inflected *is* must occupy Infl

According to (a), the auxiliary verb or copula, if it occupies Infl, must be spelled out as [Iz] (if it has the features [+3sg, +Pres]). I will call (b) the Overt Agreement Constraint. The difference between (a) and (b) is that (b) does not specify that the verb is in Infl, so that the features [be [+3sg, +Pres]] could have been formed as the result of affix hopping where Infl lowers to V, and we have the structure [$_{\text{V}}$ V Infl]. I will assume that dialects for which *bees* is not possible have the more general constraint in (b) instead of (a). On the other hand, dialects for which *bees* is possible (see (39), (40) below) have only (a). Lastly, I will assume that all dialects obey (c).

That fact that *I, you, we*, and *they* can appear as subjects with agentive *be* shows that constraints similar to (37b) do not exist for these person/number combinations. This in turn might suggest that for these person/number combinations there is no affix-hopping (see Kayne's 2000: 198). However, such an account would then lose a simple explanation for why there is do-support (in contexts of negation and inversion) for all persons and numbers (do-support is needed because affix-hopping is impossible). I propose instead that the reason agentive *be* can be used with non-third singular subjects is because these affixes are not spelled out at all after they undergo affix hopping. For example, since the features [+1sg, +Pres] are not spelled out after affix hopping (e.g., 'I run'), the unit [$_{\text{V}}$ be [+1sg, +Pres]] is not subject to phonetic spell-out. Rather, only the uninflected verb *be* is spelled out.

Similarly, the agentive *be* cannot be inflected for tense. The judgments in (38) represent my I-language, I have done no internet searches on this issue:

(38) a. *If you beed nice...
b. *You beed nice?

Therefore, agentive *be* can take no overt inflection (neither third person singular nor past tense). I call this constraint the Overt Inflection Constraint (OIC).

Partee (1977: 38) reported the following sentence produced by the child of one of her colleagues.[6] This sentence involves agentive *be*:

(39) I don't like him. He always be's mean.

---

5. Kayne (2000: Chapters 10, 11) presents several arguments that *-s* is number agreement (singular).

6. See also Matthew Dryer's FUNKNET postings, on April 5 and 8, 1997. Dryer gives an important argument that this feature of his child's speech is not the result of AAVE influence.

Partee's sentence shows that for the I-language of the child who uttered (39), the Overt Inflection Constraint does not hold (in other words, the child only has (37a) and not (37b)). In addition, in a query posted to the Linguist List (5.550, May 1994), Andrew Carstairs-McCarthy reported the following two examples of agentive *be* from New Zealand:

(40) a. So Andrew stands on his desk, and be's Alexander the Great.
b. If he be's good, he'll get to go out tonight.

To me both of the examples discussed by Carstairs-McCarthy are unacceptable, but it is fairly obvious that he is discussing the agentive *be* that is the topic of this paper. These examples from Child English and New Zealand English show that *bees* should not be ruled out in general. Rather it seems that in the colloquial, informal American English that I am describing in this paper, we must simply stipulate (37b) as a morphological constraint.

4. *Be* in the imperative and subjunctive

Clauses with agentive *be* are similar to imperatives with *be*. For example, both kinds of *be* require do-support with negation:

(41) a. Don't be too careful! (imperative)
b. *Not be too careful!
c. *Be not too careful!
d. If you don't be careful,... (agentive *be*)
e. *If you not be careful,...
f. *If you be not careful,...

Another similarity is that agentive *be* often seems more acceptable in the context of an implicit command. For example, the sentence 'If you don't be careful, you will be caught' implies the command 'Be careful!'. (Note that in both 'if you be careful' and 'if you don't be careful' the implicit command is 'Be careful!') Furthermore, the subject of an imperative clause and the subject of a clause with agentive *be* are interpreted agentively. Lastly, both agentive *be* and imperative *be* can be followed by a deleted constituent, just like *be* following a modal auxiliary (see Section 6 below):

(42) a. Don't be! (in response to the sentence 'I may be late')
b. If you be careful, you will be OK. If you don't be, you will be in trouble.
c. I won't be (in response to the command 'Don't be late!')

However, clauses involving agentive *be* are different from the imperative in that there are contexts where agentive *be* is found that do not allow null subjects (e.g., in conditionals):

(43) *If __don't be careful...

In addition, imperatives are restricted to matrix clauses, whereas agentive *be* can appear in a variety of embedded clauses. Furthermore, we saw that the past tense sequence 'didn't be' (see 21, 22, 23) occurs with agentive *be*, but such a sequence never occurs in the imperative (*'Didn't be careful!'). Lastly, while the imperative allows a verb to be in the progressive ('Don't be giving me those lame excuses', 'Be studying your Spanish when I get home!', see Akmajian et al. 1979), this does not seem to be possible with agentive *be* (see Section 6 below). Given these differences with agentive *be*, I will assume that clauses with agentive *be* are not imperatives. It may in fact be the case that the *be* used in the imperative is sometimes agentive *be*.

It is tempting to analyze clauses with agentive *be* as subjunctives, given that the subjunctive also uses uninflected *be*. Furthermore, the examples of agentive *be* often involve an implicit command, just as the subjunctive does ('I would prefer that you not do that' implies the command 'Don't do that'). Now consider the following examples (examples from Quirk et al. 1985: 1012).

(44) a. Congress has voted that the present law be maintained.
b. It is essential that a meeting be convened this week.

(45) a. If any person be found guilty, he shall have the right of appeal.
b. Whether she be right or wrong, she will have my unswerving support.
c. Though he be the president himself, he shall hear us.

Putting aside the past subjunctive ('were') and formulaic examples ('be that as it may'), the above examples represent the main categories of the subjunctive in English. The examples in (44) illustrate the mandative subjunctive which occurs as the complement of predicates expressing demand, recommendation, etc. The examples in (45) show that the subjunctive *be* can also occur in concessives and conditionals in formal English.

There are many reasons why clauses with agentive *be* cannot be analyzed as subjunctives. First, agentive *be* necessarily occurs with do-support in the presence of negation (see 1b) (compare *'if you not be careful'), but do-support is not in general possible in the subjunctive (see Potsdam 1996: 96). An example of the subjunctive without do-support is given below:

(46) The Senate has decreed that such students not be exempted from college dues. (Quirk et al. 1985: 156).

Potsdam (1996: 96) explains the lack of do-support in subjunctives by postulating a null subjunctive modal which is not an affix. The null modal occupies Infl and blocks do-support.

Second, most examples of the subjunctive in conditionals and concessives (45) are restricted to formal written English. This contrasts with agentive *be* which has a very informal colloquial flavor.

For these reasons, I will not unify the use of the formal subjunctive *be* and agentive *be*. There is a type of informal subjunctive where the use of *be* may be more closely related to agentive *be*, illustrated in the following example:

(47) a. I suggest that you don't be late next time. (informal)
b. I suggest that you not be late next time . (formal)

Sentences such as (47a) may also be related to the use of *be* with do-support in why-questions:[7]

(48) Why don't you be nice to your sister (careful, vigilant, quiet, honest, strong, like that, mean, more specific, yourself, my friend, a good boy, a model for your kids, a good Catholic parent)?

I will put aside the informal subjunctive and do-support with why-questions for further research.

In doing the internet searches, I eliminated the examples that were potentially formal subjunctives. I could most easily do this by eliminating pages that dealt with the Bible, Shakespeare (and other pages from early modern English), and Dickens.

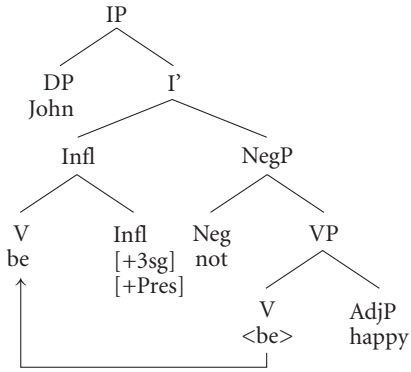
5. Agentive *be* and verb movement

I assume that the structure of the clause is the following (see Collins 1997, among many other sources):

(49) $[_{CP}$ Comp $[_{IP}$ Subj $[_{Infl'}$ Infl $[_{NegP}$ Neg vP/VP$]]]]$

VP stands for a verb without an external argument (in the case at hand the copula *be* or auxiliary verb *be*), and vP stands for a verb with an external argument. Given this structure, consider the analysis of 'John is not happy':

(50)



The verb moves to adjoin to Infl (the notation <…> indicates a copy left by movement). Then, at Spell-Out, the combination of *be*+Infl is realized as *is* (see (37a)).

7. See for example, Huddleston and Pullum (2002: 114).

Given (49), the only way that a verb can appear before the negation is if it raises to Infl. As the following sentences show, this is impossible for agentive *be*:

(51) a. If you are not careful, you will be caught.
b. If you don't be careful, you will be caught.
c. *If you ben't careful, you will be caught.
d. *If you be not careful, you will be caught.

Although the finite verb *are* can appear before negation (a), agentive *be* can never appear before negation (whether or not the negation is contracted). Similarly, agentive *be* never appears before the subject, as shown below:

(52) a. We be nice when we're trying to impress the teacher. (See 18, 20)
b. *Be we nice when we're trying to impress the teacher?

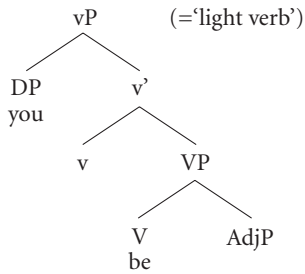
This shows that agentive *be* does not raise to Comp.

How can we explain why agentive *be* does not undergo verb movement (to Infl or to Comp)? Chomsky (1995) argues the movement of auxiliaries in English is forced by the fact that they are semantically vacuous. Chomsky's approach implies the following generalization:

(53) If a verb is not semantically vacuous, it will not be able to undergo overt raising to Infl (in English)

Putting aside the technical details of Chomsky's account, it follows from this generalization that agentive *be* should not undergo raising, and this is exactly what we find in examples such as (1b). In other words, agentive *be* is behaving like a non-auxiliary lexical verb. More specifically, I propose that the agentive component of agentive *be* is contributed by v (little v, or 'light verb') (see Chomsky 1995; Collins 1997). Therefore, the example in (1b) with agentive *be* has the following structure:

(54)



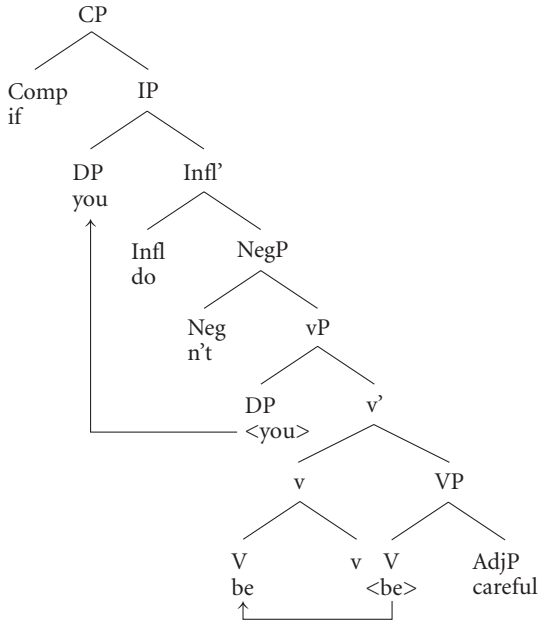
In this structure, v ('light verb') contributes the agentive meaning to the verb. The verb *be* raises and adjoins to v before Spell-Out. Since v is not semantically vacuous, it does not raise to Infl. Furthermore, I assume that V cannot move directly to Infl (by skipping over little v) because of Relativized Minimality (see Rizzi 2001).

A technical issue comes up with respect to the theta-role of the DP *you*. Assuming that in order to receive a theta-role from the adjective, the DP *you* occupies Spec AdjP

(as well as Spec vP), the question is whether there is a control or movement relation between the DP in Spec vP and the DP in Spec AdjP. I will not take up this question here for reasons of space.

The full structure for 'if you don't be careful' is the following:

(55)

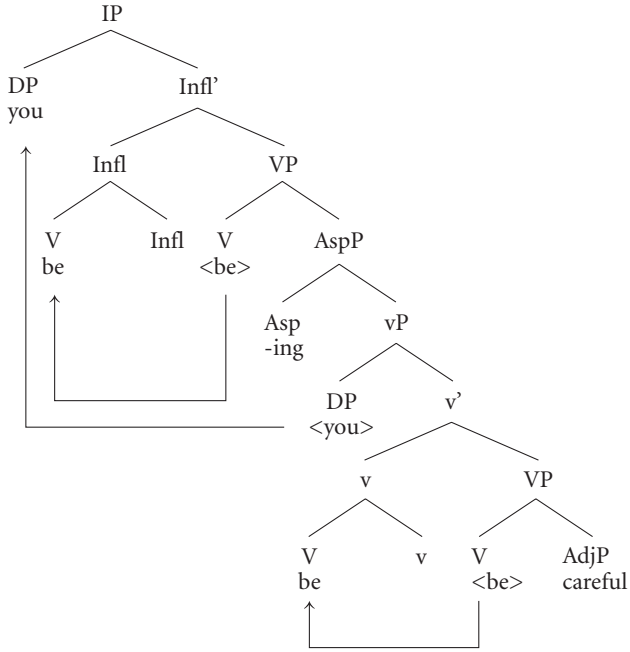


Consider now the problem of why 'if you don't be careful' must have an agentive interpretation. If vP were not generated, *be* would be forced to move to Infl (since do-support is a Last Resort operation), yielding 'If you are not careful'.

A related problem is why 'you are being careful' must have an agentive interpretation. Consider first the structure of 'you are being careful' on the agentive interpretation:[8]

8. I assume that *-ing* lowers from Asp to v in the progressive (affix hopping).

(56)



If 'you are being careful' had a non-agentive interpretation, the complement of Asp would be VP headed by *be* (not little vP). I suggest that this is impossible because the verb *be* by itself is semantically vacuous. Therefore, it is impossible for it to be the complement of Asp, which minimally needs a verb that has an event variable. In (56), little v contributes the event variable.

At the beginning of the paper (see (3)) I gave the following standard paradigm:

(57) a. John isn't happy.
b. *John doesn't be happy.
c. John shouldn't be happy.

The unacceptability of (57b) is the result of several factors. First, agentive *be* is more acceptable in the context of an implicit command (which is why it is more frequent with if-clauses and second person subjects). Nothing about (57b) suggests that it is an implicit command. Second, agentive *be* is only possible where the subject is acting agentively. Unlike *careful*, *quiet*, and *nice*, the adjective *happy* is less natural with an agentive interpretation (note the awkwardness of 'I am being happy' compared to 'I am being careful'). Since agentive *be* is not possible, the copula must raise to Infl, and do-support is ruled out.

6. Habitual *Be* in AAVE

Consider the following examples (from Green 1998: 39; see also McWhorter 1998; Wolfram & Schilling-Estes 1998 for background information and Green 2000 for a detailed semantic analysis):

(58) a. She always be a clown on Halloween.
'She always dresses as a clown on Halloween'
b. The children be at school when I get home.
'The children are usually at school when I get home'
c. He can't find his mail because it be here.
'He can't find his mail because it is always here'

As Green (1998: 56, 57) shows, habitual *be* does not undergo any type of movement.

(59) a. Becky be watching the basketball games.
b. Do Becky be watching the basketball games?
c. *Be Becky watching the basketball games?
d. Becky don't be watching the basketball games?
e. *Becky be not/ben't watching the basketball games.

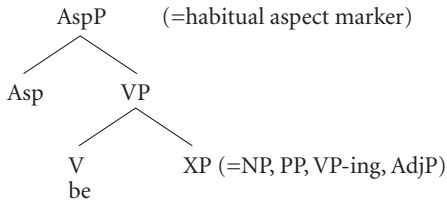
The example in (59c) shows that habitual *be* does not raise to Comp in yes-no questions. The example in (59e) shows that habitual *be* does not raise to Infl (moving to the left of negation). Similarly, Labov (1998: 120) notes that habitual *be* follows the negation and requires do-support (see (59d)).

Habitual *be* and agentive *be* share many properties. First, both agentive *be* and habitual *be* are uninflected. Second, both agentive *be* and habitual *be* require do-support in the presence of negation (unlike the formal subjunctive in English see (46)). Third, neither agentive *be* nor habitual *be* appear to the left of negation. Fourth, neither agentive *be* nor habitual *be* can appear to the left of the subject (see (59c)). Fifth, both agentive *be* and habitual *be* can occur without a preceding auxiliary (59a) (unlike the uninflected *be* in sentences like 'I will be nice'). Sixth, habitual *be* in finite clauses seems to require[9] overt subjects just like agentive *be* (and unlike the imperative).

I propose that habitual *be* heads a VP taking an XP complement, where XP can be any category. This proposal is closely related to that of Green (1998) who proposes that aspectual *be* heads a PredP, which takes a wide range of complements (VP, PP, NP, AdjP).

9. I make this claim on the basis of the examples I have seen. I have not verified this property of habitual *be*.

(60)



In this structure, habitual *be* can take a complement of any category (NP, PP, VP-ing, AdjP) (see for example, Green 2000: 19), just like (non-habitual, non-agentive) *be* in standard English. Asp (aspectual marker) contributes the habitual meaning to the verb (on which, see Green 2000).

Suppose that *be* does not raise to Asp in (60). Then by Relativized Minimality (Rizzi 2001), it will be impossible for *be* to raise to Infl (skipping over Asp). Now suppose that the verb *be* raises and adjoins to Asp before Spell-Out. Since Asp is not semantically vacuous, it does not raise to Infl (see principle (53)). In neither case can *be* raise to Infl. See the discussion following (76) for more on the issue of whether *be* raises to Asp.

Standard English does not have a null Asp head taking a VP complement for habitual interpretations ('The children are usually at school'). If it did, the Asp head should block *be* movement to Infl, yielding an uninflected *be*.

In the remainder of this section, I will discuss a number of differences between agentive *be* and habitual *be*. First, compare the structure of agentive *be* in (54), where the subject is underlyingly in Spec vP, with (60) where the subject does not occupy Spec AspP. It follows from my theory that the subject of habitual *be* does not have to be interpreted agentively. The following sentences (provided by a reviewer) show that this is the case: 'Those parties be long', 'Those papers be long'. In fact, habitual *be* can occur with subject expletives: 'If it be snowing, I'm not gon' jog'.

Another difference is that agentive *be* never appears with a following progressive participle (cf., (59a)).

(61) a. *If you be running,
b. *If you be smiling,

I propose that this constraint is related to the fact that the progressive participle of *be* cannot be followed by a progressive participle (see Akmajian et al. 1979: 19):

(62) a. You are noisy.
b. You are being noisy.
c. You are smiling.
d. *You are being smiling.
e. *You are being running.
f. *You are being giving me excuses.

Furthermore, even though *noisy* can occur with agentive *being* in the progressive (62b), the agentive *be* cannot be followed by the progressive participle *being*.

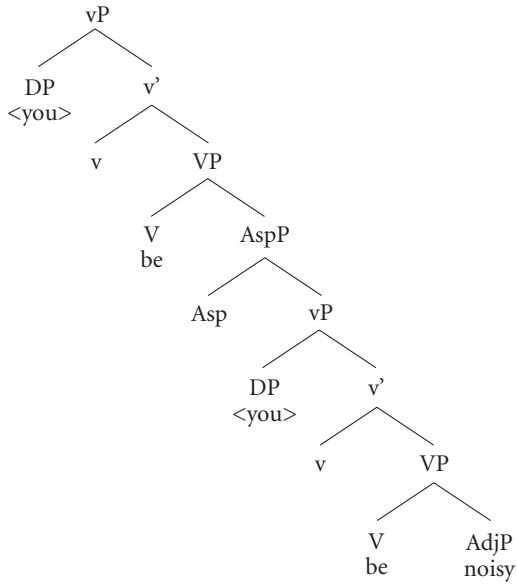
(63) a. *If you be being noisy (agentive *be*, agentive *being*)
b. *When we be being nice,
c. *Did you be being nice?
d. *How do you be being a model for your kids?

Earlier I claimed that agentive *be* raises to the head of vP (see (54)). Therefore, (63a) involves two agentive vPs (one for uninflected *be* and one for *being*, see (56)). I suggest that this sort of vP recursion is impossible:

(64) Each vP must be dominated by a distinct IP

The structure of the vP for (63a) is the following:

(65)



In this structure, there are two vPs which are not separated by an IP. If verbs like *smile* and *run* also involve vP, then the ungrammaticality of (61a, b) will be given the same explanation. Since habitual *be* in AAVE is not dominated by vP (see (60)), habitual *be* followed by V-ing in (59a) does not violate (64).

A third difference between habitual *be* and agentive *be* is the use of *bees*. The use of *bees* in AAVE is a complicated issue (see McWhorter 1998: 147 on some sociolinguistic considerations, and see Poplack & Tagliamonte 2004 for a comparison of the factors that influence the usuage of verbal -s in Samaná English and Devon English, see Kayne 2000: 207 on Newfoundland English *bees*). Some examples are given below (from Bailey & Maynor 1985: 208):

(66) a. . . .because they bees watching you when you didn't think they was.
b. . . .because they bees together all the time.
c. The cabbage bees the kind they have now.

d. In the evening, it bees hot on this porch.
e. She don't bees there.

The fact that *bees* in these examples appears with both singular and plural subjects shows that it is not a third person singular agreement suffix. Furthermore, the fact that it appears with do-support (66e) suggests that it is not an inflectional suffix (originating in Infl) at all. In fact, Labov (1998) claims that "there is no third singular /s/ in AAVE" and furthermore that "though various suggestions have been made for a functional or semantic interpretation of *bees* vs. *be*, no clear evidence to support this view has emerged" (Fn. 15). Therefore, I suggest that *bees* in AAVE does not violate the Overt Inflection Constraint, because verbal *-s* in AAVE is not inflection for person/number (see (37)).

As for habitual *be* in other English dialects, Crystal (1997:338) notes for Hiberno-English: "Copula and auxiliary *be* are used in distinctive ways, chiefly expressing contrasts of habitual action and continuity: *be* found with forms of *do* (it does be colder at nights) and also, especially in the north, with an *-s* ending (I be walking, She bees walking)." The latter sentence indicates that the Overt Inflection Constraint does not hold for one variety of habitual *be* in Irish Engish.

I will now compare an aspect of habitual and agentive *be* that appears to differ, although much more empirical work is needed. Green (1998) discusses the fact that habitual *be* cannot license a following deleted constituent:

(67) a. Bruce be singing and I do, too.
b. Bruce be singing and I be singing, too.
c. *Bruce be singing, and I be, too.

The same pattern extends to adjectival complements of habitual *be* (Lisa Green, p.c.):

(68) a. *Bruce be happy, and I be too.
b. *Bruce be nice, and I be too.
c. *Bruce be late, and I be too.

When negation is introduced, the judgments are murkier (Lisa Green, p.c.):

(69) a. *Bruce be happy, and I be too.
b. *Bruce be happy, but I don't be.
c. ???Bruce be happy, but you don't never be.

(70) *Bruce be running, but I don't be.

According to Lisa Green, (69a, b) are both bad, but (a) is worse. Furthermore, (69c) is even a little better than (69b), and may be acceptable to some people. On the other hand, (70) with V-ing, is out.

Consider now deletion following agentive *be*.[10]

10. I would like to thank Eric Potsdam for pointing out to me the difference between deletion following agentive *be* with and without do-support. In Potsdam's thesis (1998:77), he gives the

(71) a. If you be careful, you will be OK, but
if you don't be ___, you will be in trouble.
b. If you don't be careful, you'll end up in the hospital but
*if you be___, you'll be ok. (Eric Potsdam, p.c.)
c. If you don't be careful, you'll end up in the hospital but
?if you do be ___, you'll be OK. (Eric Potsdam, p.c.)

In this respect, agentive *be* is like the *be* found in imperatives, why-questions and the informal subjunctive, which also allow deletion following *be* when there is negation and do-support:

(72) a. Don't be! (E.g., 'don't be late!')
b. *Be!
c. Why don't you be? (E.g., 'why don't you be nicer to your sister?')
d. I suggest that you don't be. (E.g., 'I suggest that you don't be late')

Consider now the fact that agentive *be* does not license deletion when in the progressive (Akmajian et al. 1979: 24):

(73) a. I will be nice, and John will be __ too.
b. *I am being nice, and John is being __ too.

This fact cannot be explained by reference to agentive *be*, since clear cases of non-agentive *be* show the same constraint (see Akmajian et al. 1979: 24):

(74) a. John was being watched by the FBI, and Bill was ___ too.
b. *John was being watched by the FBI, and Bill was being __ too.

Akmajian et al. (1979) explain the possibility of deletion in (74a) by postulating an obligatory rule of Be Shift, which moves the verb *be*. The remnant VP (with a null head) can then be deleted. The analysis is shown below[11]:

(75) a. [$_{VP}$ be nice] → Be Shift
b. [$_{VP}$ be [$_{VP}$ <be> nice]] → Merge *will*, *John*
c. [$_{IP}$ John [$_{Infl'}$ will [$_{VP}$ be [$_{VP}$ <be> nice]]]] → VP deletion
d. [$_{IP}$ John [$_{Infl'}$ will [$_{VP}$ be [$_{VP}$ _____]]]]
(deletion of the remnant VP is possible)

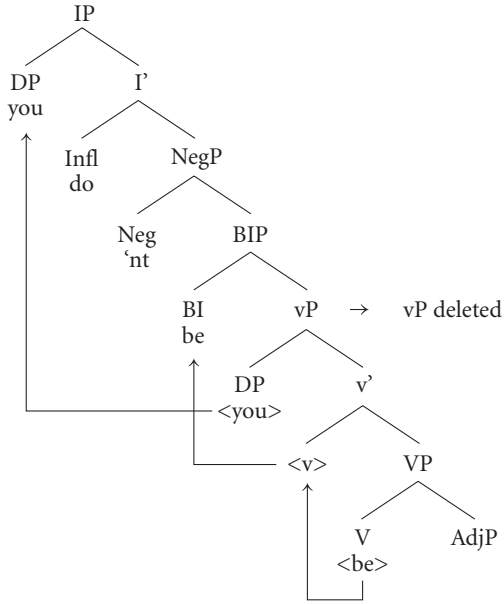
Licensing Condition on VPE: An elided VP must be c-commanded by an overt, non-affixal inflectional head within the same s-projection. In other words, if Infl is empty, deletion is not allowed. This condition explains why deletion following agentive *be* is better if there is negation with do-support. I give a different analysis below.

11. Akmajian et al. (1979: 30) claim that Be Shift is a restructuring rule.

According to Akmajian et al. (1979), the impossibility deletion following the progressive (74b) follows from the lack of Be Shift in the progressive.[12]

I will extend Akmajian et al.'s Be Shift analysis to the case of agentive *be* following do-support. Suppose that when Infl is filled with *do* (because of the presence of negation) or a modal auxiliary, it licenses an extra functional category between Infl and vP (the position of agentive *be*). This extra functional projection can be called BI, for bare infinitive. In other languages, such as French, the infinitive in an analogous position would show up with overt morphology (e.g., V-er, for the *-er* verbs in French). When BI is present, agentive *be* moves to it, and the vP complement of BI can be deleted (71a, c). This is illustrated in the following diagram (of 'you don't be'):

(76)



When BI is absent (because Infl does not contain *do*, or a modal auxiliary), v remains in-situ (occupied by *be*), and deletion following *be* is not possible (71b).

Consider now the impossibility of deletion following habitual *be* in AAVE (67–70). Suppose that *be* does not raise to Asp in AAVE (see (60)). From this it follows immediately that *be* cannot raise to BI (by Relativized Minimality). Therefore, deletion following habitual *be* would be impossible. Suppose on the other hand that *be* raises to Asp, but that Asp (as opposed to v) never raises to BI. If we made the further assumption that the VP complement of Asp cannot be deleted, it would follow that the constituent following habitual *be* could not be deleted. I leave as an open question how to distinguish these two possibilities.

12. According to Akmajian et al. (1979: 32), main verb *have* also undergoes shifting ('. . .but he will have ___ soon').

Lastly, consider the status of (69c). It may be that *be* can raise to Asp and Asp can raise to BI as a marginal option. In this case, the AspP complement of BI could be deleted. Clearly much more empirical work is needed (e.g., to what extent is Be Shift and deletion possible following other modal auxiliaries in AAVE).

The above account of deletion following agentive *be* raises the following question. Consider again (53): If a verb is not semantically vacuous, it will not be able to undergo overt raising to Infl (in English). I used this principle to explain why agentive *be*, just like main verbs (*run*, *see*) does not raise to Infl. But note that main verbs do not undergo Be Shift: *'I will seem happy and John will seem too'. Thus only *be* (and *have*) undergo Be Shift to BI, not main verbs. In other words, agentive *be* cannot raise to Infl, but it can raise to BI, and main verbs cannot raise to either Infl or BI. I have no explanation for this asymmetry between movement to Infl and movement to BI for the moment.

To summarize, agentive *be* and habitual *be* share a number of properties that can be attributed to the use of a null head, which is v in the case of agentive *be* and Asp in the case of habitual *be*. There are, however, a number of differences between the two. First, the specifier of vP, but not AsP, is a theta-position. Therefore, agentive *be* cannot be used with an expletive subject. Second, habitual *be*, but not agentive *be*, can be used with a following progressive participle. Third, habitual *be*, but not agentive *be* can take the form *bees*. Fourth, the constituent following agentive *be*, but not habitual *be*, can be deleted.

7. The source of habitual *be* in AAVE

Bailey (1982) proposed that habitual *be* in Irish English was the source of habitual *be* in AAVE. The hypothesis is discussed extensively by Rickford 1986, who rejects it. Rickford (1986: 261) points out that *be* (by itself) is not used as a habitual marker in any Carribean creoles. Rickford (1986: 266) proposes a decreolization theory of the development of habitual *be* in AAVE that takes into account the creole data, while also allowing a role for the influence of Irish English dialects (Rickford 1986: 272). In this theory, there are four stages, illustrated below (Rickford 1986: 267):

(77)	a.	He	(d)a	quiet		
	b.	He	does	quiet		(*d(a)* → *does*)
	c.	He	does	be	quiet	(*quiet* reanalyzed as adjective)
	d.	He		be	quiet	(deletion of *does*)

In stage one, a habitual marker (*d*)*a* is used in creoles. This (*d*)*a* "represents convergence between English habitual *do* and similar West African forms" (Rickford 1986: 269) (e.g., the verbal suffix -*na* in Ewe). In stage two, unstressed *does* replaces (*d*)*a*. In stage three, *be* is introduced following *does* either as a replacement for the creole copula *de*, or preceding adjectives. In stage four, *does* is deleted (due to phonological reduction processes), leaving *be* as a habitual marker.

While I think that Rickford's analysis is attractive, I would like to consider a scenario where the presence of agentive *be* played a role in the development of habitual *be* in AAVE. I suggest the following stages:

Stage one: language learners hear ambient dialects of English some of which (maybe most) contain agentive *be*. The presence of agentive *be* in informal American English, New Zealand English (see (40) in Section 3) and British English (p.c., Paul Elbourne) suggests that agentive *be* is a widespread feature of English around the world, and raises the possibility that it existed in the relevant contact dialects (Irish English dialects, British English dialects, American English dialects) during the development of AAVE.

Discussion of the geographical distribution of agentive *be* raises the question of whether there is any relationship between the presence/absence of agentive *be* and the presence/absence of habitual *be* in a dialect. For example, it is striking that informal English has agentive *be*, but not habitual *be*. Does AAVE have agentive[13] *be*? Do the dialects of Irish English that have habitual *be*, have agentive *be*? Does New Zealand English have habitual *be*? If a language had both habitual *be* and agentive *be*, a sentence with non-inflected *be* would be potentially ambiguous. Are there dialects of English with such ambiguous sentences? A much wider survey of English dialects is needed to even begin addressing this question.

Stage two: Language learners analyze the syntax of uninflected *be*. The relevant syntactic features include lack of inflection, the position following negation, negation with do-support, the presence of a subject, and lack of subject-auxiliary inversion. By virtue of Universal Grammar, they know that these properties of uninflected *be* are the result of the combination of *be* with a semantically contentful head (e.g., little v).

Stage three: Language learners adapt syntax of uninflected *be* (which they learned from exposure to agentive *be*) to express habitual aspect. Crucially, there are many possible ways to express habitual aspect (e.g., as the suffix *-na* in Ewe, or as a free morpheme (*d*)*a* for Carribean creoles), but only uninflected *be* was chosen in AAVE. The simplest hypothesis is that a language learner would keep same phonological form, and syntactic features of agentive *be*, but analyze its syntax in terms of AspP (yielding a habitual interpretation) instead of vP (yielding an agentive interpretation).

I do not think my proposal is inconsistent with Rickford's proposal. In particular, Rickford proposed that in his stage 4, habitual *does* is deleted, leaving uninflected *be*. But perhaps another way of saying the same thing is that the language learners replaced overt *does* with covert Asp. In part this replacement would have been possible because of the model that agentive *be* (involving a null head) provided of uninflected *be*.

Consider an alternative where agentive *be* played no role in the development of habitual *be*. Assume that the learner wants to express the habitual as a null head Asp.

---

13. More generally, I have not looked into imperative *be*, informal subjunctive *be* (see (47a)), agentive *be* (1) or why-question *be* (see (48)) in AAVE (see Labov 1972:71 for some examples of uninflected non-habitual *be*). All of these subjects would be extremely interesting to investigate from the point of view of my paper.

Furthermore, the learner knows that there is an uninflected *be* (from constructions like 'I will be happy', 'I want to be on time', 'Be on time!') in the ambient dialects of English. So the learner simply projects the null aspect head with the uninflected *be* (Merge(Asp, VP)) and by principles of UG we get the syntactic facts which define habitual *be*. On this scenario, there is no historical connection between habitual *be* and agentive *be*. They share the same syntax because both of them involve a functional projection with a null head (AspP and vP), that takes a VP complement (headed by *be*).

The question is what counts as sufficient evidence for the language learner to postulate the presence of a null habitual head (instead of an overt suffix, or no head at all, for example). Is it sufficient that there is no overt habitual head in the surrounding dialects of English for the language learner to motivate a null head? Or is it also necessary that they see a similar null head (the head of vP) being used with similar syntactic properties? This is a question which I am unable to answer now.

In this paper, I have not brought forth any new data on habitual *be* in AAVE. The 'fresh look at habitual *be*' referred to in the title consists in (a) pointing out its similarities with agentive *be* in informal English, and (b) raising a series of questions about the diachronic relationship between habitual *be* and agentive *be*.

References

Akmajian, A., Steele, S. M., & Wasow, T. (1979). The category AUX in Universal Grammar. *Linguistic Inquiry, 10*, 1–64.

Bailey, C.-J. N. (1982). Irish English and Caribbean Black English: Another rejoinder. *American Speech, 57*, 237–239.

Bailey, G. & Maynor, N. (1985). The present tense of be in southern black folk speech. *American Speech, 60*, 195–213.

Becker, M. (2004). *Is* isn't *be*. *Lingua, 114*, 399–418.

Chomsky, N. (1995). *The Minimalist Program*. Cambridge MA: The MIT Press.

Collins, C. (1997). *Local Economy*. Cambridge MA: The MIT Press.

Crystal, D. (1997). *The English Language*. Cambridge: CUP.

Green, L. (1998). Aspect and predicate phrases in African-American vernacular English. In S. Mufwene et al. (Eds.), *African American English* (pp. 37–68). London: Routledge.

Green, L. (2000). Aspectual be-type constructions and coercion in African American English. *Natural Language Semantics, 8*, 1–25.

Huddleston, R. & Pullum, G. K. (2002). *The Cambridge Grammar of the English Language.* Cambridge: CUP.

Kayne, R. (2000). *Parameters and Universals*. Oxford: OUP.

Labov, W. (1972). *Language in the Inner City*. Philadelphia PA: University of Pennsylvania Press.

Labov, W. (1998). Co-existent systems in African-American vernacular English. In S. Mufwene et al. (Eds.), *African American Vernacular English* (pp. 110–153). London: Routledge.

McWhorter, J. (1998). *The Word on the Street*. New York NY: Plenum Trade.

Partee, B. (1977). John is easy to please. In A. Zampolli (Ed.), *Linguistic Structures Processing* (pp. 281–312). New York NY: North Holland.

Pollock, J.-Y. (1989). Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry, 20*(3), 365–424.

Poplack, S. & Tagliamonte, S. (2004). Back to the present: Verbal *-s* in the (African American) English diaspora. In R. Hickey (Ed.), *The Legacy of Colonial English: Studies in transported dialects* (pp. 203–223). Cambridge: CUP.

Potsdam, E. (1996). *Syntactic Issues in the English Imperative.* Ph.D dissertation, Santa Cruz, University of California.

Quirk, R., Greenbaum, S., Leech, G., & Svartvick, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Rickford, J. R. (1986). Social contact and linguistic diffusion: Hiberno-English and New World Black English. *Language, 62*, 245–289.

Rizzi, L. (2001). Relativized minimality effects. In M. Baltin & C. Collins (Eds.), *The Handbook of Contemporary Syntactic Theory* (pp. 89–110). Oxford: Blackwell.

Wolfram, W. & Schilling-Estes, N. (1998). *American English*. Oxford: Blackwell.

Oral narrative and tense in urban Bahamian Creole English[1]

Stephanie Hackert
University of Regensburg, Germany

This paper presents a quantitative analysis of past inflection in urban Bahamian Creole English. The variable has held a prominent position in the study of African American Vernacular English and Caribbean creoles, but whereas phonological, grammatical, and social factors have received much attention in this context, the influence of style or discourse type has rarely been noted. I analyze the formal and functional properties of three types of oral narrative obtained in sociolinguistic interviews. All of them, personal narrative as well as folktale and 'generic' narrative, show considerably lower inflection rates compared to the 'chat' mode but do so for different reasons; this, in turn, has repercussions for cross-variety comparisons, as samples usually differ in terms of their composition with regard to discourse type.

1. Introduction

As a linguistic variable, past inflection has received wide attention in the study of African American Vernacular English (AAVE) and Caribbean English-lexicon creoles (CECs). At least part of this attention is owed to the variable's central position in the enduring debate about the contemporary status of AAVE and its origins and relations with CECs (cf. Poplack & Tagliamonte 2001: 103). Whereas the absence of the dental suffix on regular verbs like *work* was first interpreted as a result of the more general phonological process of word-final consonant cluster reduction (Labov et al. 1968; Wolfram 1969; Fasold 1972), Bickerton (1975: 159) maintained that phonological factors "merely serve the [...] purpose of masking" the effects of grammatical factors, such as stativity, which, in his 'typical' creole tense-mood-aspect system (e.g., Bickerton 1981: 58), is inextricably linked with the presence or absence of tense marking.

---

1. An earlier version of this paper appeared in Hackert (2004). Thank you to two anonymous reviewers for this volume for their detailed and valuable comments.

Since the 1970s, analyses of past inflection in AAVE and CECs have become more and more comprehensive; they now often investigate not only phonological and social factors but also include various grammatical distinctions, such as verb category, lexical and grammatical aspect, temporal relationship and/or disambiguation, etc. (cf., e.g., Singler 1984; Rickford 1986; Winford 1992; Weldon 1996; Blake 1997; Patrick 1999; Poplack & Tagliamonte 2001; Hackert 2004). What has not received comparable amounts of attention is intra-speaker variation according to style or discourse type.

This paper investigates the influence of various types of narrative discourse on the occurrence of past inflection in urban Bahamian Creole English (BahCE), a CEC with close historical links with AAVE. Marking rates vary considerably between non-narrative speech, narratives of personal experience, folktales, and generic narratives; the following sections consider each narrative type separately in order to uncover topical, functional, and/or grammatical constraints affecting overt morphological marking of past temporal reference.

2. Data

The data that form the basis of this study were collected during three visits to the Bahamian capital, Nassau, in 1997 and 1998, each of them between one and three months long.[2] All narratives occurred during sociolinguistic interviews (cf. Labov 1984: 32–42), which covered topics such as work, traditional crafts, family life, life on the 'Out Islands'[3] of earlier times, individual episodes in Bahamian history, or folklore. I conducted sociolinguistic interviews with over thirty speakers, twenty of which – balanced for the parameters of age and sex – were chosen for detailed analysis. My interviewees were between 25 and 81 years old; all of them stemmed from working-class areas of Nassau (cf. Hackert 2004: 21–24).

Table 1 displays the various formal ways of talking about the past in urban BahCE and their distribution in the output of three selected speakers.[4] It shows that even

2. My research was made possible through the generous support of the *Deutscher Akademischer Austauschdienst* and the University of Heidelberg's *Graduiertenkolleg "Dynamik von Substandardvarietäten."*

3. In the years since World War II, the Bahamas have undergone a drastic urbanization process. Today, roughly two thirds of the just under 300,000 Bahamians (CIA 2003) live in the capital, Nassau (New Providence); the only other major city, Freeport (Grand Bahama), has roughly 40,000 inhabitants. The remainder of the population is scattered among sparsely populated and less developed 'Out' or 'Family' Islands.

4. Table 1 is restricted to proportional relationships. Possible realizations (e.g., modals as *might*, *would*, *coulda*, etc.) or marking properties (i.e., whether verbs are overtly marked for past temporal reference or not) are not taken into account. While copula and auxiliary *be* are, of course, different functionally, they are lumped together here in accordance with current practice

Table 1. Distribution of past-reference structures for three speakers

	Jeanne		Albert		Carol	
	N	%	N	%	N	%
unmarked/inflected lexical verb *Yeah, that's fifteen years ago when my daddy* ***die****. I was in Miami, and I* ***had to*** *leave* – (Sidney 6:4–5)[5]	593	68	211	51	211	54
be **copula/auxiliary** *This* ***Ø*** *between seven – eight o'clock in the night. And I didn't know they's watching, and I pull – I had a few dollars and I* ***was*** – (Henry 4/3/97, 4:38–40)	177	20	64	16	91	23
been *I* ***been*** *to Andros, yeah, couple of Friday 'go.* (Mr. Jones 6:1)	1		9	2	25	6
used to *You been here like when I* ***used to*** *had my car one time, I did take you-all and drive myself.* (Mr. Jones 3:28–29)	27	3	47	11	7	2
preverbal markers (*did*, *done*) *But my – my – my husband days – he never sponge. Uh-uh. The sponge* ***did done*** *dead.* (Mrs. King 16:28–29)	4		22	5	20	5
modal *And then all I* ***coulda*** *do is just sit down and look at the number and try to keep it in my head though – though I was in pain.* (Carol 4/15/97, 10:21–22)	2		10	2	3	1
(Past) Perfect *have* *But other than that, life* ***has*** *been going well for me.* (Sister Brown 36:26)	10	1	–		1	
do*-support** *What* ***did *daddy bring?* (Mrs. Miller 5/14/97, 8:4–6)	2		–		–	
verbal *-s* *And his sister was there, and she* ***says****, Oh* – (Mrs. Smith 1:49)	1		–		–	
negative structures *But they say someone did die right in our house. I* ***ain't never*** *see – nothing in here.* (Shanae 15:7–8)	52	6	47	11	36	9
Total	**869**	**100**	**410**	**100**	**394**	**100**

in creole and AAVE studies, where investigations of variable copula absence generally include auxiliary *be*. All forms of copula or auxiliary *be* (i.e., zero, contracted, or full) were included. With regard to the other two 'primary' verbs (Quirk et al. 1985:96), *have* and *do*, uses in auxiliary function (e.g., *do*-support, perfect *have*) or as preverbal particles have been separated from main-verbal ones. Negative structures are listed separately but are not differentiated, as they may be seen as a field of investigation in its own right (cf. Schneider 1997). Inconsistencies in the percentage columns are owed to the fact that percentage values below one are not given.

5. All speaker names are pseudonyms; the numbers refer to page and line of the transcript. If multiple recordings of a single speaker entered into the data base, a recording date occurs.

though past temporal reference in BahCE involves preverbal particles and a variety of auxiliary structures apart from variably inflected lexical verbs, the latter account for between roughly half and two thirds of all past-reference structures. Unmarked verbs, in turn, are more frequent than inflected ones; the overall past-inflection rate of the sample is 32%.

The following quantitative analyses are based on a total of 8,172 tokens of the variable (ed), i.e., unmarked or inflected past-reference lexical verbs. Excluded were copula and auxiliary structures (except those involving the 'semi-auxiliary' *have to*; cf. Quirk et al. 1985: 137) as well as structures involving modal verbs or preverbal particles. Also discounted were all structures in which a past event time is not identical to reference time (generally rendered in standard English by means of the present perfect or the past perfect); included, however, were past habitual situations. With the exception of the devoicing category (e.g., *send*), for which token numbers were too small to make generalizations, all of the lexical and morphological verb categories commonly distinguished in quantitative analyses of past inflection were included in the tabulations. Verbal *-s*, even if in a past context, was discounted (cf. Hackert 2004: 118–138).

As in other CECs and in AAVE, the occurrence of past inflection in urban BahCE depends on various linguistic and extralinguistic factors. One of the strongest constraints is verb category (cf. Winford 1992: 322; Weldon 1996: 65; Blake 1997: 154; Patrick 1999: 235; Poplack & Tagliamonte 2001: 118), with non-syllabic regular verbs (e.g., *play*, *work*) strongly disfavoring inflection (with marking rates of 16% and 6% respectively) and individual high-frequency lexical items like *have* or *go* exerting a highly favorable influence (86% and 68% marking). A verb which occurs particularly frequently in narrative contexts is *say*; as in other varieties, this verb occurs predominantly in unmarked form (16% inflection).

3. Oral narrative and tense

Most generally, narratives revolve around "someone telling someone else that something happened" (Fleischman 1990: 100). Within this broad category, three subtypes are represented in my sample:

– The story is a specific past-time narrative that has a plot and makes a point (Fleischman 1990: 102). Most stories of the present sample fall under the heading of "narratives of personal experience" (Labov 1984: 32), in which narrators tell of a memorable incident out of their own lives (or, less commonly, the life of individuals familiar to them).
– Folktales are a part of the cultural heritage of a people; they narrate the experiences of particular fictional, often animal, characters. The most popular of these characters in the Bahamas are B'Rabby, an underdog and trickster, and his friend B'Bouky.

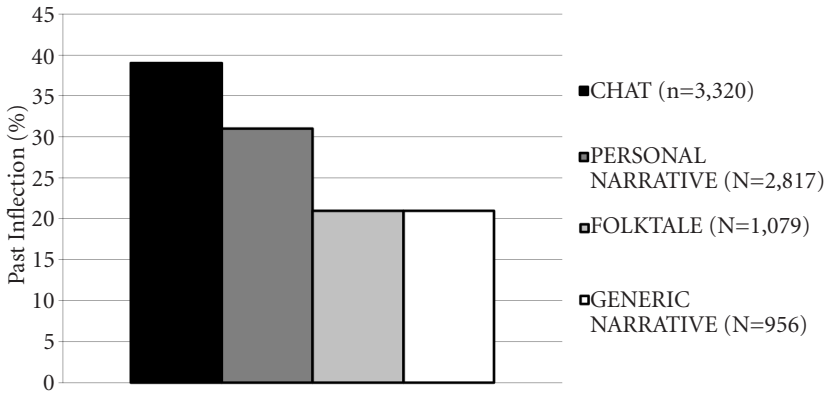


Figure 1. Past inflection according to discourse type

– Generic narratives "relate what used to be the case in the past [...] and are marked for the feature [+habitual]" (Fleischman 1990: 104).

Non-narrative speech is subsumed under a 'chat' mode. The chat mode is the default discourse type of this analysis. It is not bound by any particular topic or structure, and utterances were assigned to it after all the other discourse types had taken their share of the pie.

Figure 1 shows rates of past inflection according to discourse type. The use of past inflection varies considerably, the chat mode being marked almost twice as frequently (39%) as the folktale (21%) and the generic narrative (21%). The narrative of personal experience falls in between (31%).

The question is what motivates this variation. The following sections consider the different narrative types as well as their formal and functional properties individually.

3.1 Narratives of personal experience

A subclass of the story, the narrative of personal experience, has received wide attention in sociolinguistics. Within the framework proposed by Labov and Waletzky (1967), this type of narrative is defined as a recapitulation of past experience in which the events are presented iconically, i.e., in the order in which they happened (cf. also Labov 1997). The narrative of personal experience can be divided into smaller units with regular functional and formal properties; there are complicating action (CA) clauses, which present the story events one by one and thus propel the narrative forward; orientation clauses, which provide the listener with the appropriate background; and evaluation clauses, which communicate the narrator's feelings about and attitudes toward the events – the point of the story. An abstract may precede the narrative and encapsulate the events to come or present the general proposition of the story. The

coda, finally, returns narrator and listener to speech time. This is illustrated in the following narrative (Sidney 6: 4–20).

I – I been in a – a – in – in a – in a – plane crash before. [No way!]	ABSTRACT
Yeah, that's 15 years ago when my daddy die. I was in Miami, and I had to leave – out of Miami about – 7:30 that night. And catch a private plane out for Crooked Island, Sunday morning for the funeral.	ORIENTATION
And when we got – when we got in Crooked Island, before we – just before we land, the pilot said – tell us to be prepare,	COMPLICATING ACTION
'cause the gear stuck, the landing gear stuck. But, you see,	EVALUATION
Crooked Island is a place like this, where the e- where – where the runway is – both side is sandy ground, and the plane –	ORIENTATION
when it hit – when it hit the runway, it run off the – the runway. And she hit and she run until she bog in the sand – just about –	COMPLICATING ACTION
just about the top of the plane was in the sand. But other places wouldn'ta had that opportunity, 'cause 'e ain't much places in the Bahamas where the runway is sandy. And that what save us.	EVALUATION
And – and – and – that will be a good idea if the Bahamas government could set up runway like that to save – to save life. And that's how – our life was – was a nine-seater – that's how our life was save.	CODA

Table 2 displays the distribution of verb forms and their marking properties in 25 such narratives from 19 different speakers. As for forms of *be* in copula or auxiliary use, both contracted and zero forms were counted as unmarked for temporal reference, whereas *was*, *were*, and *been* were considered marked.[6] *Used to*, *did*, and *done* are invariant with respect to their marking properties, with *used to* and *did* indicating habitual and relative past respectively, *done* designating completive aspect (cf. Winford 1993: 31–66). Excluded from the analysis were all verb structures that did not unambiguously refer to the past, i.e., present- or inclusive-time and hypothetical constructions, structures pertaining to situations outside of the story world, as well as quoted speech, direct

6. This was done because zero and contracted forms of *be* are not restricted to present contexts but also occur with past temporal reference in BahCE. The same holds for *is*. Shilling (1978: 64) gives the following example: "The first Bahamian captain ever captain a boat from Europe out here əz a Bah- was a Ragged Islander". According to her, "there seems to be no a priori reason for /əz/ in past contexts not to represent *was* with deleted initial consonant." Even though, as illustrated in Table 1, *been* may also occur in utterances that resemble standard English sentences like "My sister has been to Rome" (Quirk et al. 1985: 194), it clearly functions as a past copula in contexts such as "She been Adventist for twenty-odd years" (Sister Brown 8: 5).

Table 2. Verb forms and their marking properties across narrative units

	Abstract		Orientation		CA		Evaluation		Coda	
	N	%	N	%	N	%	N	%	N	%
***be* copula/auxiliary**	4	50	128	73	5	80	42	81	9	78
preverbal *did*			16				1		1	
preverbal *done*			5							
used to			22							
modal			5	20			16	56		
***do*-support**	2	100	12	83			22	91	3	100
unmarked/inflected lexical verb										
– *have*			31	97	13	100			1	100
– *go*	1	100	6	100	23	100	3	100	2	0
– *say*	1	0	3	0	126	7	3	33	2	50
– regular non-syllabic	5	40	21	0	105	0	11	9	5	0
– regular syllabic			7	14	27	11	4	25	2	50
– irregular	13	62	36	22	201	18	18	44	8	0
Total	**26**	**58**	**292**	**64**	**500**	**18**	**120**	**65**	**33**	**42**

or indirect.[7] *Have*, *go*, and *say* are listed separately on account of their frequency and idiosyncratic marking behavior (cf. above).

Table 2 shows that past-reference *be*, *have*, or *used to* cluster in orientation clauses, which communicate to the listener the time and place of the events recounted as well as the identity and motives of the characters involved. These structures may be viewed as semantically stative[8] and tend to be marked for past temporal reference. CA clauses, in contrast, which narrate the events of the story in iconic form, contain almost exclusively lexical verbs, roughly four fifths of which are uninflected. Note that the likelihood of such verbs to be inflected increases slightly if they occur in an orientation rather than in a CA clause. The makeup of evaluation clauses is more mixed, with *be* and lexical verbs occurring in almost equal proportions, but both marked even more frequently than in orientation clauses.

Why, however, should narratives of personal experience foster the absence of past inflection if compared to the chat mode? One approach would be to focus on the role of

7. Of the structures listed in Table 1, negative structures and *have* auxiliary structures (N=1) also do not appear in Table 2.

8. Even though statives, progressives, and habituals differ in various ways and therefore do not form a single aspectual category (statives are a type of lexical, or situation, aspect; progressives and habituals are two imperfective viewpoints, or grammatical aspects), they all present verb situations as open, i.e., without endpoints, and thus have similar functions in (narrative) discourse. It is true that progressives and habituals are usually non-stative at the level of the underlying individual verb situation and therefore retain the syntactic properties of this verb situation; semantically and functionally, however, they closely resemble "proper" statives (Smith 1997: 50–51, 84–86).

narratives of personal experience within the sociolinguistic interview. Such narratives, and specifically those triggered by the 'danger of death' question, had originally been intended by Labov as a solution to the problem of eliciting the vernacular, the idea being that they constituted one of the contexts in which casual speech occurs (cf. Labov 1984: 33). The attention to content necessary in the recounting of personally significant or even life-threatening incidents was thought to lower attention to speech; less attention to speech would then result in more casual speech and thus allow linguists to gain access to the vernacular.

Another interpretation of the variation between marked and unmarked verb forms in creole narratives (e.g., Bollée 1977; Corne 1977) has been in terms of a historical present (HP) analysis. According to this interpretation, the unmarked forms occurring in the CA sections of narratives would serve to describe the past "as if it is happening now", conveying "something of the dramatic immediacy of an eye-witness account" (Quirk et al. 1985: 181). One obvious fact militating against an HP analysis for creoles is the sheer preponderance of zero-marked forms: roughly four fifths of the 495 lexical verbs occurring in the CA sections of the 19 narratives discussed here occur in their unmarked form; in the English narratives investigated by Schiffrin (1981: 51), only 30% of the verbs in CA clauses are in the HP.

It does not appear necessary, however, to appeal to Labovian speech styles or a (somewhat vaguely defined and controversial) category like the HP (cf., e.g., Wolfson 1979; Leith 1995) to account for the low rates of past marking in creole narratives of personal experience. An alternative explanation may be sought in the fact that the vast majority of tokens of (ed), i.e., of unmarked or inflected lexical verbs, that occur in such narratives occur in CA clauses (cf. Table 2), where they usually instantiate the perfective (PFV), an aspectual category which is often unmarked crosslinguistically (Bybee et al. 1994: 90–91). According to Dahl (1985: 78), a 'prototypical' PFV can be characterized as follows: "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."

This definition makes two important points: perfective verb situations typically have past temporal reference, and they are presented as bounded with initial and final endpoints (Smith 1997: 66). The boundedness of perfective verb situations has an important implication for their use in discourse: the PFV may function in those clauses which recount a series of events in the order in which they occurred in the (real or any imagined) world, i.e., narrative CA clauses. In fact, "[t]he expression of a sequence of actions is one of the most characteristic functions of perfective verbs in an extended context" (Forsyth 1970: 9–10, cit. in Smith 1997: 92).

What locates events in sequence in narrative discourse is the movement of reference time: each (main clause) event in a narrative sets up a reference time, the 'current' reference time (Couper-Kuhlen 1989: 19), which is then relevant for situating the next event into the discourse. As the PFV presents events with their initial and final endpoints, the reference time of the following event cannot be that of the current one; rather, it is a new reference time immediately posterior to the one obtaining, and sequentiality arises. In sequenced clauses, contextual disambiguation is at its maxi-

mum; therefore, overt morphological marking indicating temporal reference becomes dispensable (cf. Comrie 1985: 61).

The automatic movement of reference time in the CA sections of narratives of personal experience may be both what underlies the lower rates of past inflection in this discourse type compared to the chat mode in urban BahCE and what permits the absence of an overt marker of past temporal reference in the case of the English HP. As for the former, while tokens of (ed) in the chat mode may also be perfective in aspect, they are not necessarily sequenced, but (as pointed out above) sequenced CA clauses comprise the bulk of all tokens of (ed) occuring in narratives of personal experience; it is thus not surprising that rates of past inflection are considerably lower in the latter than in the former.

With regard to the parallel between unmarked verbs in creole narratives and the English HP, Patrick (1999: 221) points to "a principle of wider application – namely, that the need for past-marking is reduced where organization of the discourse makes it redundant." The structural makeup of English and creole narratives of personal experience is, of course, the same: sequentiality is the defining characteristic of this type of narrative (Labov & Waletzky 1967: 20) in any language or variety. Where English and creoles differ is that tense-marking is categorical in the former but variable in the latter; the tightly circumscribed context of personal narratives, however, allows for the relaxation of this principle in English and even lower rates of past marking than usual in creoles. Thus, "an explanation on discourse principles [...] accounts for both data sets in a general way. [...] where syntactic variability makes space for it, discourse functions will govern the distribution of grammatical elements across linguistic contexts" (Patrick 1999: 221).

A question that remains is why individual perfective, sequenced CA verbs should be past-marked at all, i.e., whence the 18% inflection occurring at the bottom of Table 2? It has been pointed out above that, as in other CECs and in AAVE, past inflection in urban BahCE is subject to various linguistic and extralinguistic constraints, one of the most important being verb category. Individual high-frequency lexical items also show particular marking propensities. This holds not only for *have*, *go*, and *say*, which have long been known for their idiosyncratic behavior (Bickerton 1975: 104); it also affects past inflection patterns across morphological verb categories and underlies the apparent effect that stativity has on past marking (Patrick 1999: 259; Poplack & Tagliamonte 2001: 142; Hackert 2004: 160, 164–165). *Want*, for example, is marked at a considerably higher rate than other syllabic verbs in BahCE; at the same time, it accounts for a large part not only of syllabic verbs but also – together with *have* and *think* – of stative verb situations.

Another factor favoring past inflection in urban BahCE is temporal disambiguation via adverbial. While a general [±adverbial] distinction has no effect whatsoever (Hackert 2004: 178), the presence of a durative adverbial (Quirk et al. 1985: 481–482), such as *up to 1960*, clearly raises inflection rates for perfective verb situations (Hackert 2004: 179). Finally, as would be expected, extralinguistic factors, such as social class or education, also affect the use of past marking.

To sum up, the lower rates of past inflection observed for personal narratives vis-à-vis the chat mode in urban BahCE may be explained with the fact that the vast majority of tokens of (ed), i.e., of unmarked or inflected lexical verbs, in the former occur in sequenced, contextually disambiguated perfective CA clauses. But why should folktales and generic narratives show even lower marking rates than the narrative of personal experience?

3.2 Folktales

At first sight, storytelling in the Bahamas seems to be a dying pastime. First, only eight of the twenty speakers of my sample told folktales at all, with two of them producing no more than minimal tales containing five and 17 tokens of the variable (ed) respectively. Second, only two of the more prolific storytellers of the sample were under fifty. Comments from the Bahamian public and the literature seem to corroborate this picture. When asking around for good storytellers, I was often told that younger Bahamians no longer know folktales because they watch too much television. Bahamian folklorist and writer Patricia Glinton-Meicholas also bemoans the demise of the 'old-story':

> Unfortunately, few people 'talk ol' story' or even remember one nowadays. Like many 'people intensive' activities, storytelling has fallen prey to modernization. Electronic and other forms of entertainment are anaesthetizing the creativity of our oral storytellers and the need which produced them. (Glinton 1994: 12)

Nevertheless, the fate of the traditional Bahamian folktale may not be as grim as it seems. In the first place, even in the 1950s, when American anthropologist Daniel Crowley was recording folktales in the Bahamas, a particular speaker's ability and willingness to tell folktales seems to have been more a matter of attitude than of age:

> It must be made clear that not all Bahamians tell stories. A few refer to them as 'lies' and 'trash', and explain that their parents tried to make them 'logical' and would not teach them such nonsense. (Crowley 1966: 13)

Perhaps more importantly, the years since independence from Britain (1973) have seen a dramatic increase in interest in things Bahamian; particularly educated Bahamians seem to be reappraising local traditions in religion, music, and literature and to be uncovering their African roots, and a number of folktales have been collected and published in this process (e.g., Glinton 1994; Maynard 1994).

All 'old-story', as folktales are called in the Bahamas, comprise "a set of formulaic characters having specific names, functions, and personality characteristics" (Crowley 1966: 28). The most popular of these characters are B'Rabby, an underdog and trickster, and B'Bouky, his greedy, gullible friend. Bahamian 'old-story' often resemble theatrical productions in that the narrator makes characteristic use of facial expressions, gestures, and voice (a 'sperrit', for example, always speaks nasally). They occasionally include a brief song (or 'sing'); traditionally, there is also a formulaic opening and closing. The following story illustrates.

> I – I heard it from – no, my grandparents and other people, you know. Uh-huh. They would – they would start it, Once upon a time was a merry good time. The monkey chew tobacco, and he spit white lime.
> [That's how they started it?]
> Cockaroach jump from bank to bank, and he never touch water. Then they'll say – this was Bo- B'Bouky and B'Rabby. Say they were – they were two brothers. Say now they – they – they – they went to this place where these spirit live. You heard 'bout spirits?
> [Yeah – oh, can you tell me that?]
> Say then – say then they – they – the – the spirit were living up in the sky – they's – their house been up in the skies there – and – and then when they go to the – the spirit-them leave, and they go out. Bouky and B'Rabby would go to this house, say now – the – the – when they – when they got there, they – they sing this song for the house to come down – say the house came tumbling down to the ground – say they went inside the house, and they cook, and they eat all of these food, and B'Bouky was the greedy one. Say so after they finish eating they te- B'Rabby say to B'Bouky, Time to go, my bulla! B'Bou- B'Bouky – B'Bouky say, Not yet. I wants get some more to eat. Say when they listen, they hear the spirit come home. Say, boy, that time, say, the – the – B'Bouky and B'Rabby run under the spirit bed. Say then B'Bouky is so greedy, say the spirit come feed their children on the floor – in front of the bed where B'Bouky was. Say instead of B'Bouky don't keep quiet, he tell the – the spirit little girl, Gimme some, gimme some, gimme some! Say – say and – child keep a-putting little – putting piece in his hand and say he keep on begging, Gimme piece, gimme piece, gimme piece! Say the spirit little girl start hollering, say – say – say and after she started to cry, she get up and she stretch out herself, she say, Me got two daddy. Say – spirit say, What two daddy? I'm your daddy. I is your daddy. Say and – yeah! Say, Say that again! Say, I got two daddy. Say – say, What? Say, Where the next one is? Say, One up there to the table and one under the bed and he eye shine like a dollar. The spirit say, Wait! Say, Lemme get – lemme find him!
> [And then what happened?]
> Say – the – the – the spirit – when the spirit look under there and he see B'Bouky, he pull him out, say and he hot some water, and he scold him up and say he throw him – he pitch him out the sky down on the ground and say whe- when B'Bouky (…) – when he fall on the ground, he burst wide open. Then after that, she say, Be, bo, ben, the story is end. And let go of the pullets and to catch the hen.
> (Mrs. Miller 4/1/97, 1:1–27)

Returning to the role of past marking in folktales, it was shown in Figure 1 that the inflection rate of this discourse type is considerably lower than that of narratives of personal experience and amounts to only about half the rate evidenced in non-narrative speech. Why should this be so? In principle, what has been said for narratives of personal experience holds for folktales as well: they are stories, i.e., specific past-time narratives with a plot and a point, and can be divided up into abstract, orientation, complicating action, evaluation, and coda. The important difference is that, whereas

narratives of personal experience recount unique events out of the lives of real people, folktales narrate the adventures of imaginary characters, be they human or animal. The difference, in other words, is that between fact and fiction.

According to Fleischman, fiction can be recognized not only in "the text's pragmatic context (i.e., the participants in the narrative contract and their relationship to the utterances of the text)" (1990: 107), but it also has linguistic correlates. As for a text's pragmatic structure, according to Fleischman, the prime characteristic of fiction is the narrator's act of "transferring 'origin,' of attributing the text to someone else", and of creating an embedded communicative context (1990: 107). This transfer of origin can be observed in the above tale as well. Both explicitly (*I heard it from – no, my grandparents and other people*) and with her frequent use of *say*, which, when not accompanied by an overt grammatical subject and a change in notional subject, does not denote an actual event of saying but merely signals the continuation of quoted matter, the narrator, Mrs. Miller, emphasizes that she is not the original producer of the story but only retells what she has heard from others.

A second characteristic of fictional narrative is grammatical. Both fiction and nonfiction, according to Fleischman (1990: 110), have as their basic, unmarked tense the PAST. In non-fictional discourse, the PAST has the task of relegating the events narrated to an anterior time sphere. If the events recounted are not real, however, how can the the PAST assign them past temporal reference? According to Fleischman, in fictional narration the PAST loses its original referential value and acquires an expressive function instead, suggesting verisimilitude by endowing a fictional discourse with a historical "as if:" "let us pretend for the duration of this textual transaction that such and such events occurred in some past world" (1990: 112). If a morphological mark of past temporal reference loses its referential function, however, it becomes dispensable – which is precisely what happens in the folktales described above. This holds only in part for the orientation clauses of these tales; note that in Mrs. Miller's tale information about the characters involved (*they were two brothers*) and the location of the events to follow (*the spirit were living up in the sky*) is overtly past-marked. It might be argued that, in order to retain the appearance of verisimilitude and locate the story in the – whatever distant and/or unspecific – past, the employment of at least some past morphology is necessary. In CA clauses, however, past inflection is extremely rare.

To sum up, the occurrence of past inflection in folktales is subject to two related but distinct constraints: first, just as in narratives of personal experience, the bulk of all tokens of unmarked or inflected lexical verbs occurs in CA clauses, which are sequenced and therefore highly restricted in terms of temporal reference; second, as the PAST need not signal actual past temporal reference, it has expressive rather than purely referential function and can therefore be abandoned more easily than in nonfictional texts. A third factor must be considered. It has often been pointed out that the characteristic feature of folktales is their lack of originality. Folktales are handed down from generation to generation, unchanged in basic content and structure. The purpose of storytelling is often seen in the display of a society's basic values and beliefs; at least in the Bahamas; however, according to Crowley (1966: 12), stories are told just

as often "to entertain and to gain prestige through entertaining". Both aspects require the possibility to actuate the events told on any given occasion; the resulting narratives elude precise location in time (as well as in space) but take their significance from their timeless functions; the predominance of forms unmarked for temporal reference reflects this.

3.3 Generic narratives

As indicated above, generic narratives are marked for the feature [+habitual]; they narrate not a specific past incident but relate what used to be the case in the past or what usually occurs in the present (Fleischman 1990: 104). As such, they are more akin to lists than to stories (Schiffrin 1994: 293–330) – a resemblance which has repercussions for the instantiation and meaning of various kinds of information structure.

Most basically, whereas stories tell about something that happened to someone, lists describe a category by presenting a collection of items that are the same in one respect (justifying their inclusion in the category constituting the list) but different in other regards (laying the foundation for the segmentation of the list into subcategories). Whereas the basic unit of the story is an event that moves reference time forward, the list's basic unit is simply an 'entity', i.e., "anything of which something may be predicated (individuals, sets, propositions [...])" (Schiffrin 1994: 297). However, such entities may also be constituted by events – which is what endows accounts of past habitual situations with a narrative-like character.

Apart from being a skilled teller of folktales, Mrs. Miller was also one of the most prolific producers of generic narratives in my sample, as the following excerpt illustrates.

> My grandmother, she farmed – the island when my granddaddy and my dad went out to sea, you know – so like –
> [What did she farm? What were some of the –]
> Oh, corn, sweet potatoes, beans, peas, cabbages, toma-
> [Cassava, too?]
> Oh, cassava, we love cassava. We have boiled cassava and crab. You – you heard of crab?
> [Yeah, but I've never seen it. I want to go to Andros –]
> That'll be nice. Yeah. We cook – we cook conch and rice, and conch and grits, and salt beef – you know of salt beef? Uh-huh, and we cook with fresh toma- we never use no tomato paste and things like that. Fresh (...) – we – we raise these things, and we grow these things – uh-huh. And we – we – we cook with them. Sometime, we have the conch dry – we bruise it out and we hang it on the line and let it –
> [The conch – you hang the conch on the line?]
> Yeah. You – you – you bruise it out, and then you string it together, two together. You let that dry, and you cook some good pea soup with that. With pea soup and dumpling with some salt beef in it. And you would – you would – you would lick your fingers off. Yeah. Then – then we cook – we cook peas and rice with coconut

> in it. Coconut milk. And that used to be so good. And then you have your fry fish on the side. We raise our own chicken. We – when we want to cook chicken, we kill a chicken, and we – we – cook that – when we want meat, we kill it, and then you – some used to wring the neck like that, and – and then you pick it off, and you clean it up. When we want meat, we don't have no meat from Nassau, we – we call it town, no boat coming from town – to town. That's used to call it. Sometime, you know, I be telling stories, we ha- our church women, we have – uh – plays, and they always call on me and another lady to tell – to tell things what happen when we were growing up on the Family Island, you know.
>
> (Mrs. Miller 4/16/97, 4: 34–5: 10)

One of the most significant differences between stories and lists concerns the role played by descriptive statements. As indicated above, in stories, description usually occurs in orientation clauses, where it has a backgrounding function. If it is embedded within a story's CA section, it may also be evaluative (Labov & Waletzky 1967: 32). In lists, description plays a very different role. It is the defining characteristic of this discourse type; thus, rather than being endowed with a backgrounding function, description constitutes the "skeletal organization" of a list (Schiffrin 1994: 293). Description builds on stative predicates; not surprisingly, structures involving "the most central and characteristic of stative verbs", *be* and *have* (Quirk et al. 1985: 205), are the most frequent verbs in the orientation sections of stories (cf. Table 2).

At first sight, generic narratives, involving events rather than states, seem to contradict this requirement. All of the verb situations contained in Mrs. Miller's account of food preparation in times past, however, are marked for the feature [+habitual]; some (*that used to be so good* and *some used to wring the neck like that*), in fact, are accompanied by an overt morphological mark of (past) habituality. Habitual verb situations, in turn, are semantically stative – they present situations as open, i.e., without endpoints (Smith 1997: 50). Thus, even though generic narratives present a number of events (which is what makes them similar to stories and justifies their inclusion in the narrative category in the first place), there is no temporal progression toward a resolution. Rather, events in generic narratives are descriptive; they represent situations "that can serve as continuing instantiations of a given category" (Schiffrin 1994: 302).

It is true that generic narratives often contain sequenced events; in fact, in chicken killing, one can't clean the chicken, then pick its feathers, and then wring its neck.[9] Such sequences are often part of process descriptions. Note, however, that the process description just reproduced realizes only a subcategory to the second of the two main categories constituting Mrs. Miller's list, i.e., foodstuffs and their preparation (Hackert 2004: 198–199). Whereas the temporal structure of process descriptions – just like that of stories – is prepositional in that there is an iconic sequence of events, the temporal structure of lists is exclusively textual: it is concerned with the relations between the en-

9. As one reviewer rightly pointed out.

Table 3. Past inflection by grammatical aspect

	N	%	factor weight
perfective	6,404	35	.61
habitual	1,746	21	.16
Total/p_i	8,150	32	.27

tities presented successively in the list. The use of *then* neatly illustrates this difference (cf. Schiffrin 1994: 300–301); whereas in

> when we want to cook chicken, we kill a chicken, and we – we – cook that – when we want meat, we kill it, and then you – some used to wring the neck like that, and – and then you pick it off, and you clean it up

the adverb indicates progression through reference time, in

> then we cook – we cook peas and rice with coconut in it. Coconut milk. And that used to be so good. And then you have your fry fish on the side

it simply introduces two Bahamian side dishes, i.e., it indicates progression through text sequence. In contrast to stories, where the sequence of events cannot be altered without changing our interpretation of what happened, the order of entities that have the same categorial status (such as two different side dishes) in a list has little relevance to the structure of that list.

As a factor influencing past marking in CECs and AAVE, habituality has never received the amount of attention that, e.g., stativity has – in spite of Bickerton's early observation (1975: 150) that in Guyanese Creole verb situations denoting habituality strongly inhibit past inflection. An exception is Winford (1992: 335), who finds confirmation for this statement in his AAVE and Trinidadian Creole data. Poplack and Tagliamonte (2001: 144) identify the same pattern in varieties of Early African American English. Not surprisingly, as Table 3 shows, grammatical aspect, i.e., the perfective/habitual distinction, exerts a strong influence on past inflection in urban BahCE as well,[10] as shown particularly by the Varbrul factor weights.[11]

The strength of this influence becomes even more impressive through a look at the two most frequently marked lexical items, *have* and *go*. Table 4 shows that both marking rates and factor weights drop drastically if only habitual verb situations are considered.

---

10. There are 22 unmarked or past-inflected progressive tokens, e.g., "And all them doctors getting me scared now. I so afraid – I wonder what's was happening" (Sister Brown 4: 11–12), that have been excluded from Table 3.

11. Varbrul is a statistical computer package that establishes probabilities of application for a variable linguistic rule, such as past inflection (cf., e.g., Young & Bayley 1996). Factor weights above .5 indicate favoring effects, those below .5 disfavoring ones.

Table 4. Past inflection by grammatical aspect, *have* and *go* only

	have			*go*		
	N	%	factor weight	N	%	factor weight
perfective	515	98	.77	343/ 365	94	.82
habitual	347	69	.15	21/ 170	12	.04
Total/p_i	862	86	.93	364/ 535	68	.78

Why should habituality be so favorable to the absence of marking? Even though habitual verb situations "differ from generic ones by their lack of lawlikeness" (Dahl 1985: 97), the two categories are very closely related (Dahl 1985: 95); generic sentences, in turn, are frequently expressed via "the most unmarked TMA cateogry, as in English, where the Simple Present is used" (Dahl 1985: 100). Thus, it does not seem uncommon crosslinguistically for habitual and/or generic verb situations to lack overt morphological marking; the disfavoring effect that habituality has on past inflection in urban BahCE and other CECs and AAVE may reflect a universal rather than a creole-specific trend.[12]

To sum up, while generic narratives disfavor past inflection in urban BahCE as strongly as folktales do, they do so for very different reasons. As stories, folktales consist largely of sequenced and thus temporally disambiguated verb situations; the effect that such situations have on variable past inflection is compounded by their occurrence in a fictional context and the timeless functions fulfilled by such tales. Generic narratives, on the other hand, instantiate habitual situations; overt marking of such situations is frequently absent, in creoles as well as in other languages.

4. Conclusion

The preceding sections have shown that discourse type, instantiated via various types of oral narrative, considerably affects past inflection in urban BahCE. Compared to non-narrative speech, i.e., the chat mode, all narrative types show lower marking rates, but they do so for different reasons.

This finding has repercussions for the comparability of samples across varieties. While the elicitation of narratives of personal experience is one of the explicit goals of the sociolinguistic interview (Labov 1984: 32), which forms the basis of all quantitative studies of linguistic variation, the telling of folktales is not always part of this speech

12. An alternative, phonological explanation is forwarded by Poplack and Tagliamonte, who find that in Early African American English "[s]tem forms are clearly favored in contexts where *would* + verb is an option in (Standard) English" (2001: 149) and suggest that "contamination of the preterite data with tokens of deleted *would* – which are indistinguishable from zero preterites – may well be responsible for the favoring effect of habitual aspect on zero realization of past temporal reference" (2001: 150).

event, and at least some studies of past inflection in AAVE and CECs (e.g., Winford 1992: 318; Blake 1997: 150) explicitly exclude habitual tokens as well. This means that overall rates of past inflection will be higher in such samples than in samples like the one used for this study. If we are to arrive at reliable comparisons between various CECs on the one hand and between AAVE and CECs on the other, we must take such differences into account, i.e., we must spell out the composition of samples in terms of discourse type or, better even, compare rates of past marking for the discourse types individually.

References

Bickerton, D. (1975). *Dynamics of a Creole System*. Cambridge: CUP.

Bickerton, D. (1981). *Roots of Language*. Ann Arbor MI: Karoma.

Blake, R. (1997). All o' we is one? Race, Class, and Language in a Barbados Community. Ph.D dissertation, Stanford University.

Bollée, A. (1977). *Le créole français des Seychelles*. Tübingen: Niemeyer.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, aspect, and modality in the languages of the world*. Chicago IL: University of Chicago Press.

CIA (2003). *The World Factbook*. http://www.cia.gov/cia/publications/factbook/print/bf.html.

Comrie, B. (1985). *Tense*. Cambridge: CUP.

Corne, C. (1977). *Seychelles Creole Grammar: Elements for Indian Ocean proto-creole reconstruction*. Tübingen: Narr.

Couper-Kuhlen, E. (1989). Foregrounding and temporal relations in narrative discourse. In Alfred Schopf (Ed.), *Essays on Tensing in English. Vol. 2: Time, Text and Modality* (pp. 7–29). Tübingen: Niemeyer.

Crowley, D. (1966). *I Could Talk Old-Story Good: Creativity in Bahamian folklore*. Berkeley/Los Angeles: University of California Press.

Dahl, Ö. (1985). *Tense and Aspect Systems*. Oxford: Blackwell.

Fasold, R. (1972). *Tense Marking in Black English: A linguistic and social analysis*. Washington DC: Center for Applied Linguistics.

Fleischman, S. (1990). *Tense and Narrativity: From medieval performance to modern fiction*. Austin TX: University of Texas Press.

Forsyth, J. (1970). *A Grammar of Aspect: Usage and meaning in the Russian verb*. Cambridge: CUP.

Glinton, P. (1994). *An Evening in Guanima: A treasury of folktales from the Bahamas* (2nd edition). Nassau: Guanima Press.

Hackert, S. (2004). *Urban Bahamian Creole: System and variation* [Varieties of English around the World G32]. Amsterdam: John Benjamins.

Labov, W. (1984). Field methods of the project on linguistic change and variation. In J. Baugh & J. Sherzer (Eds.), *Language in Use: Readings in sociolinguistics* (pp. 28–53). Englewood Cliffs NJ: Prentice-Hall.

Labov, W. (1997). Some further steps in narrative analysis. *Journal of Narrative and Life History, 7*, 395–415.

Labov, W. & Waletzky, J. (1967). Narrative analysis: Oral versions of personal experience. In J. Helm (Ed.), *Essays on the Verbal and Visual Arts: Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society* (pp. 12–44). Seattle WA: University of Washington Press.

Labov, W., Cohen, P., Robins, C., & Lewis, J. (1968). *A Study of the Non-standard English of Negro and Puerto Rican Speakers in New York City*. New York NY: Columbia University.

Leith, D. (1995). Tense variation as a performance feature in a Scottish folktale. *Language in Society, 24*, 53–77.

Maynard, P. D. (1994). *B'Booky and B'Rabby: Stories and folktales of the Bahamas*. Nassau: Bahama Vision.

Patrick, P. L. (1999). *Urban Jamaican Creole: Variation in the Mesolect* [Varieties of English around the World G17]. Amsterdam: John Benjamins.

Poplack, S. & Tagliamonte, S. (2001). *African American English in the Diaspora*. Oxford: Blackwell.

Quirk, R., Greenbaum. S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.

Rickford, J. R. (1986). Past marking in the Guyanese mesolect: A close look at Bonnette. In K. M. Denning, S. Inkelas, F. C. McNair-Knox, & J. R. Rickford (Eds.), *Variation in Language: NWAV-XV at Stanford* (pp. 379–394). Stanford CA: Stanford University Department of Linguistics.

Schiffrin, D. (1981). Tense variation in narrative. *Language, 57*, 45–62.

Schiffrin, D. (1994). *Approaches to Discourse*. Oxford: Blackwell.

Schneider, E. W. (1997). The cline of creoleness in negation patterns of Caribbean English creoles. In R. Hickey & S. Puppel (Eds.), *Language History and Linguistic Modelling: A festschrift for Jacek Fisiak on his 60th birthday. Vol. 1: Language History* (pp. 1055–1067). Berlin: Mouton de Gruyter.

Shilling, A. W. (1978). Some Non-standard Features of Bahamian Dialect Syntax. Ph.D dissertation, University of Hawai'i.

Singler, J. (1984). Variation in Tense-aspect-modality in Liberian English. Ph.D dissertation, University of California at Los Angeles.

Smith, C. S. (1997). *The Parameter of Aspect* (2nd edition). Dordrecht: Kluwer.

Weldon, T. L. (1996). Past marking in Gullah. In M. Meyerhoff (Ed.), *University of Pennsylvania Working Papers in Linguistics, 3* (pp. 63–72). Philadelphia PA: University of Pennsylvania.

Winford, D. (1992). Back to the past: The BEV/creole connection revisited. *Language Variation and Change, 4*, 311–357.

Winford, D. (1993). *Predication in Caribbean English Creoles*. Amsterdam: John Benjamins.

Wolfram, W. (1969). *A Sociolinguistic Description of Detroit Negro Speech*. Washington DC: Center for Applied Linguistics.

Wolfson, N. (1979). The conversational historical present alternation. *Language, 55*, 168–182.

Young, R. & Bayley, R. (1996). VARBRUL analysis for second language acquisition research. In R. Bayley & D.S. Preston (Eds.), *Second Language Acquisition and Linguistic Variation* (pp. 253–306). Amsterdam: John Benjamins.

Aspects of variation in educated Nigerian Pidgin

Verbal structures

Dagmar Deuber
University of Freiburg, Germany

This paper presents quantitative analyses of variation in selected areas of the grammar of Nigerian Pidgin as spoken by educated speakers (who also have a good command of English, the lexifier): tense/aspect marking, copula forms and verbal negation. The background is the theory going back to work by DeCamp and Bickerton in the 1970s on Jamaica and Guyana, respectively, that a linguistic continuum inevitably develops in certain situations of contact between a pidgin/creole and its lexifier. The data presented here, however, do not indicate the existence of a pidgin-to-English continuum in Nigeria. This may be taken as evidence that the Caribbean creole continua have arisen due to specific sociohistorical circumstances and that such continua are not bound to develop elsewhere.

1. Introduction

Nigerian Pidgin (NigP) is an English-based expanded pidgin which fulfils important communicative and integrating functions in Nigeria's multilingual and multiethnic society, whose linguistic repertoire also includes English, the *de facto* official language, and about 500 indigenous languages (Grimes 2000). Once a mere trade language, NigP is now used as a lingua franca in a wide range of contexts (Agheyisi 1984: 212), and it has been described as "the most widely spoken language in the country" (Faraclas 1996: 2). It is mostly spoken as a second language but creolization is under way (cf. Shnukal & Marchese 1983; Faraclas 1986, 1987). While private communication remains NigP's most typical domain of use, government agencies have, in the past few decades, increasingly exploited its obvious usefulness as a medium of information transmission to the public (Agheyisi 1984: 12; Elugbe 1995: 294), notwithstanding the fact that explicit official recognition has still not been extended to the language. This development has been particularly pronounced in the mass media, especially the radio as the medium with the broadest audience (see e.g. Mann 1998; Deuber 2002; Deu-

ber & Oloko 2003). A further aspect of the expansion of NigP is that while it used to be disdained as an inferior language by the educated, who also have English at their disposal as a lingua franca, younger educated Nigerians, especially university students, have increasingly adopted it as an informal medium of communication and a means of expressing of solidarity (Agheyisi 1984: 212). As Mann (1993: 171f.) observes,

> Nigerian University campuses have become a primary locale for communication in A.N.P. [Anglo-Nigerian Pidgin]. And it is not enough to say, in this case, that the ethnic mixture of the student population is wholly responsible for this state of affairs. A.N.P. has become a trendy code for the young generation, in which they discuss their lectures, social life, politics, and love affairs. It serves as a marker of the new detribalized generation of Nigerians. It is not unusual to find two or more students of the same ethnic origin – who can communicate equally well in their mother tongue – resorting to A.N.P for intragroup communication.

The expansion of NigP into new domains of use and more diverse segments of society has received considerable attention from scholars interested in the status and functions of the language, but it is not reflected to the same degree in studies of structural aspects of NigP such as its grammar, which is the concern of the present study. Some limited analyses of social variation have been conducted (e.g. Agheyisi 1984; Faraclas 1986), and selected grammatical aspects of educated NigP use in the diaspora have been investigated (Poplack & Tagliamonte 1996; Tagliamonte, Poplack & Eze 1997). However, the major contributions to the field such as, in particular, Faraclas (1996) have continued a tendency going back to the earliest studies (e.g. Agheyisi 1971) to focus on the "purest form" (Agheyisi 1984: 214) of the language; this is usually associated with less educated speakers whose English is too limited for it to exert a major influence on their NigP use. By contrast, the educated variety, whose spread raises important questions such as that of the nature and degree of the influence of the lexifier - which, after all, might be thought to threaten the existence of NigP as an independent linguistic system – has not been a major topic. The present paper is intended as a contribution towards redressing this imbalance in studies of NigP grammar. It looks at educated NigP as used in private, informal communication as well as radio broadcasts. Verbal structures have been chosen for a detailed analysis of grammatical variation. The paper's major question is whether or not a pidgin-to-English continuum obtains in Nigeria. Three aspects of verbal structures are examined with this question in mind: tense/aspect marking, copula forms and verbal negation. The basis for the continuum question will be explained in Section 2. Then, after an introduction to the data and method of the study (Section 3), selected text samples and the results of the analyses of verbal structures in the corpus will be presented and discussed (Sections 4–6). The final Section (7) will draw some conclusions regarding the question of a NigP-to-English continuum.

2. The question of a NigP-to-English continuum

Work on language contact situations of the type investigated here, i.e. between a pidgin/creole and its lexifier, used to be dominated by the view that given a certain social mobility, a linguistic continuum will inevitably develop. Thus, through a process of 'depidginization' or 'decreolization', the original pidgin/creole, in the continuum the basilect, and the lexifier or acrolect become linked by a spectrum of new intermediate varieties referred to as the mesolect, whose users tend to be more educated and socially mobile than speakers of the basilect, and urban rather than rural residents. Sebba (1997: 210f.) illustrates the kind of linguistic variation that is typical of creole continua by the following Jamaican sample sentences, which represent versions of the same sentence ranging from the basilectal to the acrolectal one:

me a nyam – me a eat – me eatin' – I eatin' – I is eatin' – I am eating

The view that a continuum develops in situations of continued contact between a pidgin/creole and its lexifier combined with certain social conditions is associated in particular with DeCamp's and Bickerton's work in the 1970s on the situation in Jamaica and Guyana, respectively (DeCamp 1971; Bickerton 1973, 1975b). From these cases, the idea was extended to others. For example, in a paper on Tok Pisin presented at a conference in 1973 – two years before Papua New Guinea's independence – (published as Bickerton 1975a), Bickerton argued that once Papua New Guinea became independent and there was an increase in social mobility, a continuum would probably develop between Tok Pisin and English. As parallel cases supporting this prediction he cited not only the Guyanese one but also that of Nigeria, claiming that there were no varieties intermediate between NigP and English until the country's independence in 1960 because the social conditions favouring such a development were not present, but that since then a continuum had emerged. Todd (1974: 51f.) even went so far as to say that pidgin/creole continua can be found in all areas of the world where an English-based pidgin/creole coexists with English. However, neither this sweeping generalization nor Bickerton's specific assertion that a continuum had developed in Nigeria were supported by the kind of detailed empirical analysis on which for example Bickerton's work on the Guyanese continuum is based. Furthermore, Agheyisi, one of the foremost Nigerian linguists who have worked on NigP, has explicitly denied the existence of a pidgin-to-English continuum in Nigeria. In her study of social varieties of NigP, she concludes:

> That NPE [Nigerian Pidgin English], like pidgins and creoles spoken in other communities, is a dynamic system is amply borne out by the co-existence in the same community of the above varieties of the language. However, it is important to point out that the relationship which exists between these varieties and NSE [Nigerian Standard English] cannot be described as that of a continuum, in the same sense that creole-speaking communities in the Caribbean region [...] have been shown to manifest this phenomenon. (Agheyisi 1984: 229f.)

Although it contradicts Bickerton's and Todd's views cited above, Agheyisi's statement, based on a selective analysis of NigP speech samples, did not spark a controversy, nor has the issue been followed up in further studies, as scholars have rather continued to focus on what would be the basilect in a continuum. Whether a mesolect comparable to Caribbean ones has developed in the NigP use of speakers with significant exposure to English is thus a question which remains to be addressed on the basis of further empirical investigation. In the present paper, verbal structures have been singled out for an analysis intended as a first step in this direction because they are an area in which variation in the 'prototypical' continua of Jamaica and Guyana is well documented, so that these cases can be used as a point of reference. In particular, reference will be made to Bickerton's analysis (by means of implicational scales) of copula forms (1973: 648ff.) and the verb phrase (1975b) across the Guyanese continuum, and Patrick's (1999) quantitative study of variation in the Jamaican mesolect, which includes a detailed analysis of past marking (Chapters 6–7).

3. Data and method

The corpus on which the present study is based consists of transcriptions of NigP speech amounting to a total of about 80,000 words.[1] The speakers in the corpus were all residents of the metropolis of Lagos at the time the data were collected and they have generally completed secondary education or higher. In an area where a pidgin/creole continuum exists, such speakers would of course not normally be expected to use the basilect, at least not regularly in spontaneous speech – it also has to be taken into account that speakers usually command a greater or lesser span of the continuum and can therefore modify their speech to a certain extent according to the context. The range of contexts covered by the corpus is shown in Table 1.[2] The range is such that if there was a NigP-to-English continuum, some of the texts, e.g. the conversations among the educated informants, would almost certainly be in the mid-to-upper mesolectal range, whereas a somewhat lower level of the continuum might by represented in some of the broadcasts (which are aimed at the segment of the audience which cannot follow broadcasts in English).

---

1. The corpus was compiled during a six-month field trip undertaken by the author in 2000. I gratefully acknowledge the financial support extended to me during that period by the German Academic Exchange Service (DAAD). I would also like to express my gratitude to Prof. Abiodun Adetugbo, then Head of the English Department, University of Lagos, for supporting the project, as well as to numerous members of this department and particularly Dr. Patrick Oloko for their contributions to the research. My sincere thanks are further due to the research assistants who helped in collecting and transcribing the data.

2. More detailed information on the corpus and the different types of texts included is given in Deuber (2002: 197ff.).

Table 1. The corpus (* Each text consists of approximately 2000 words)

Text categories		Number of texts*	Text codes
Face-to-face speech:			
	Interviews	5	V01-05
	Discussions	5	D01-05
	Conversations	10	C01-10
Radio broadcasts:			
	News	5	N01-05
	Advice	5	A01-05
	Drama	10	R01-10

Table 2. Transcription symbols

/	Intonation unit boundary suggesting nonfinality
//	Intonation unit boundary suggesting finality (as at end of declarative sentence)
?	Intonation unit boundary suggesting question
–	Truncated sentence/intonation unit
text& &text	Sentence/intonation unit continued (when interrupted)
#	Truncated word
(text)	False start
{text} {text}	Overlapping speech
[?text?]	Uncertain transcription
[?numberWORDS?]	Unintelligible speech (approximate number of words)
[+text+]	Non-NigP speech
[DESCRIPTION]	Noise/action

The texts in the corpus are transcribed in modified English orthography. The reasons for the choice of this type of transcription and the main strategies employed in adapting the English orthographic system are explained in Deuber (2002: 200f.). Transcription symbols which occur in the citations from the corpus in the following sections of the study are explained in Table 2.

The study is in essence quantitative but a discourse-analytic perspective is integrated to some extent into the analyses.

4. Analysis of text samples

Below an extract from one of the conversations in the corpus is given as a first text sample. The conversation was recorded at the University of Lagos in May 2000. The two speakers, who belong to different southern Nigerian ethnic groups (speaker A: Urhobo; speaker B: Efik), were then enrolled at the university as undergraduate students. Both are male and were in their twenties at the time the recording was made.

Text sample 1 (*References to the corpus consist of the text code (see Table 1) and the line numbers)

	Original version	Translation
	B: [...] you go see some kind lecturers wey be say/ dem go dey give you some kind lectures wey– imagine/ dem go	**B:** [...] You'll see lecturers who'll give you lectures where– imagine, they'll give you a course outline, but to
5	give you course outline but to lecture you/ zero// and dis thing na mathematics o//	lecture you, zero. And this thing is mathematics.
	A: hm//	**A:** Hm.
	B: just imagine when dem go	**B:** Just imagine, when they'll
10	tell you say/ dis na di topic wey we dey enter/ dem go just represent am by A B C//	tell you, this is the topic we're entering, they'll just represent it by A B C.
	A: ah//	**A:** Ah.
	B: for examination na im dem	**B:** Then in the examination
15	go come give you real values// [LAUGHTER] wetin you go come do? [LAUGHTER]	they'll give you real values. So what are you going to do then?
	A: man student go confuse for exam hall o//	**A:** Man, students will get confused in the exam hall.
20	**B:** ah di thing so surprise me last semester//	**B:** Ah, the thing so surprised me last semester.
	A: hm?	**A:** Hm?
	B: just take example dis quadratic equation/ just	**B:** Just take for example this quadratic equation, just
25	imagine make your lecturer just go put quadratic formula for board&	imagine your lecturer putting a quadratic formula on the board&
	A: {shoo//}	**A:** [exclamation]
	B: &{come} represent am by	**B:** &and representing it by X I
30	X I Y Z//	Y Z.
	A: e no tell una anything?	**A:** He/she didn't tell you anything?
	B: just– e no tell you anything/ e just comot// at least if e for put real digits now/ man for	**B:** Just– he/she didn't tell you anything, he/she just left. At least if he/she had put real
35	{understand}&	digits one would've understood&
	A: {understand//}	**A:** Understood.
	B: &small about am/ but man/	**B:** &a little of it, but man,
	A: na wa o//	**A:** [exclamation]
	B: people just dey carry F	**B:** people just got F's.
40	anyhow now// (C04 13–36)*	

A striking feature of this text sample are the many specialized English lexical items that one would not find in a conversation between NigP speakers who do not have this level of education. Many of these are nouns, e.g. *lecturers* (l. 2), *lectures* (l. 4), *quadratic equation* (l. 24); there also some verbs of this type, e.g. *represent* (ll. 12, 29). This phenomenon is of course not confined to language contact situations involving a pidgin/creole and its lexifier but commonplace in any situation where local languages coexist with an exoglossic official language in a broadly diglossic relationship. Consider, for instance, text sample 2, which has been taken over from Agheyisi's study of "language interlarding in the speech of Nigerians" (1977). The text sample is an extract from a conversation between two students in which the basic language of interaction is Yoruba, one of the three major ones among Nigeria's indigenous languages (the other two being Igbo and Hausa).

Text sample 2

	Original version	Translation
	A: Kinní morpheme yẹn … ǹtorí problem yẹn, ó prove difficult, sort of. Ṣẹ́ẹ̀rí ǹkan tí mo try lati ask nisẹyin …;	**A:** That thing about morpheme … because that problem, it proved difficult, sort of. You know, something that I tried to ask previously …;
5	ṣẹ́ẹ̀rí, iru 'oní-kinní kán kán' yẹn ṣé 'ó' yẹn is a separate morpheme?	you know that 'oní-kinní kán kán', is that 'ó' (is) a separate morpheme?
	B: Kì í ṣe 'ó' … 'ó' kò jẹ separate morpheme nigbati o	**B:** It is not "ó" … That "ó" is not a separate morpheme when
10	ba sọọrọ agentive. 'Oní-' naa l' o wa ninu 'alaṣọ' yẹn. T' o ba jẹ pe noun l' o pari ẹ l' ọ́un, k' o ti mọ wipe 'oní-' yẹn l' o maa separate; nitoripe 'oní-' is	you are talking about agentive. It is the 'oní-' that is in that 'alaṣọ'. If it is a noun that ends it over there, then you will know that it is the 'oní-' that
15	always joining with … noun l' o maa ri 'oní-' mọ́ in all cases. (Agheyisi 1977: 102f.; underlining in the original)	you will separate; because 'oní-' is always joining with … it is with a noun that you will always find 'oní-' in all cases.

Here, the same phenomenon can be observed, i.e. the speaker's Yoruba is interlarded with English lexical items, in this case mainly linguistic terminology such as *morpheme* (l. 1), *agentive* (l. 10), *noun* (l. 12). In addition to these items, there are some larger discourse chunks in English (e.g. *is a separate morpheme*, ll. 6–7). Such discourse chunks in English occur in the NigP corpus as well. A problem is that because of the lesser linguistic distance between NigP and English they can be more difficult to isolate than in the case of a language like Yoruba. For the present investigation, the whole corpus, split up into a large number of short extracts, was subjected to a questionnaire-based analysis by Nigerian informants competent in both NigP and English; these were asked

to identify English elements in the extracts presented to them (in a selective rather than exhaustive manner, though, which means that the results indicate only tendencies).[3] This yielded a total of almost 900 different items. About two thirds of these are single lexical or grammatical items, the latter including English verb forms as described in Section 5 below. The remainder are idiomatic expressions and discourse chunks of varying size, up to a whole clause or more. Those of the larger English segments selected by the informants which stand out most clearly according to fixed criteria – namely, that they consist of a whole clause or more and are separated from the surrounding discourse by intonation unit boundaries and/or by virtue of being a quotation or similar extraneous element – were subsequently taken account of in the transcription by applying the symbol for 'non-NigP speech' and changing the orthographic conventions to standard English ones. This is illustrated by text samples 3a–d. These are all taken from interviews and conversations, as it is mainly in the texts in the face-to-face speech section of the corpus that such segments occur, although there are also some instances in radio broadcasts (mostly in the radio drama texts; cf. also Deuber & Oloko 2003).[4]

Text samples 3 (a)–(d)

		Original versions	Translations
a		our people get proverb say/ na person wey dey stay for house na im dey do wetin? na im dey sweep di house// oyibo people dem put am in anoder way/ dem say '[+as you make your bed/ so you lie on it+]'// (V01 14–17; speaker: postgraduate student, female, age group 20–29, Urhobo ethnic group)	Our people have a proverb which says that he who lives in a house, it is he who does what? It is he who sweeps the house. The white people put it in another way, they say: . . .
b		how person (go) go cook for house/ e go pack all di dirty for house go throway for anoder man land/ because say dem never build house for dere// if house full dat	How can somebody cook in his house, and then clear away all the rubbish in the house and go and throw it onto somebody else's land because no house has been

3. For further details of the questionnaire study of corpus extracts see Deuber (2002) and Deuber and Oloko (2003).

4. In the radio texts, which are scripted and in which punctuation as in written language has been used, punctuation marks were taken as the equivalents of intonation unit boundaries for the purpose of identifying elements to be marked as non-NigP speech.

Text samples 3 (a)–(d)

		Original versions	Translations
		street where dem dey put am? for gutter// if rain fall dem go turn di so-called	built there? If there are houses everywhere in that street where do they put it?
	10	drums where dem dey put dirty inside gutter/ water go carry am// okay/ make I give di best solution for refuse disposal for Lagos// make I	Into the gutter. When it rains they'll empty the so-called drums where they put the refuse into the gutter, and the water will
	15	speak small English// [+there should be designated specific areas for refuse disposals// [. . .]+] (V02 65–72; speaker: postgraduate student, male, age group 30-39, Edo ethnic group)	take it away. Okay, let me give the best solution for refuse disposal in Lagos. Let me speak English for a moment. . . .
c		I dey listen to one person for television di oder day// im talk about marginal-	I listened to somebody on television the other day. He talked about marginalization. He said: . . .
	5	ization// e say '[+Biafra is not the solution to uh–+]' (C09 184–186; speaker: undergraduate student, male, age group 20–29, Igbo ethnic group)	
d		e be like say for dat last (lec#) semester na only dat woman be say im own lecture e dey [?1WORD?]	I think that last semester it was only that woman whose lecture wasn't– people didn't try to enjoy the
	5	e no– people no dey try to enjoy di lecture// if you know you come early come (en#) listen to am/ you no go enjoy am// but oders wey	lecture. If you knew and you came early to listen to it you wouldn't enjoy it. But others who were making an effort, people– . . . That one wasn't
	10	(dey) dey try sha people– [+we had this Englishman too// that one was not teaching anything//+] dat one no dey teach// e no dey	teaching. He wasn't teaching.
	15	teach// (C06 123–128; speaker: undergraduate student, male, age group 16–19, Yoruba ethnic group)	

Here, the contrast between NigP and English segments is not only obvious to anybody who is familiar with NigP to some degree but in two cases (3a and 3b) it is also made explicit by the speaker by means of metalinguistic commentary. Furthermore, the English segment in 3a is a proverbial expression inserted as a whole into the NigP discourse while in the remaining cases the English segments illustrate well-known discourse functions of code-switching: In 3b, the switch to English occurs after the speaker has given an account in NigP of personal experiences with the problem of refuse in Lagos; in the statement in English he then suggests a general solution to the problem. This corresponds to the code-switching function that Gumperz (1982: 80) describes as "personalization versus objectivization". Text sample 3c contains an instance of code-switching to mark a quotation (cf. Gumperz 1982: 75f.). In 3d the switch may have been triggered by the associations conjured up by the topic of the Englishman but "reiteration" (Gumperz 1982: 78) or repetition of a message in a different code, which, according to Gumperz (ibid.) often occurs as an emphatic device, can also be observed in this example.

The corpus data which will be presented in the following section will make it clear how important it is that instances of code-switching like the ones described above are identified and taken into account in quantitative analyses of the corpus.

5. Results of corpus analyses

5.1 Tense/aspect marking

As amply illustrated by the text samples in the preceding section, verbs in NigP are not inflected, but they can be combined with a set of preverbal tense/aspect markers, which includes the following: *dey* for imperfective aspect (progressive and habitual), *don* for perfective aspect, *been* for past tense, and *go* for future tense (cf. Agheyisi 1971: 133ff.; Faraclas 1996: 195ff.). (*Go* may be more appropriately characterized as a marker of irrealis modality, but since, like in English the modal *will*, it serves as the primary means of expression of future time, it is described by Agheyisi (1971: 134) and Faraclas (1996: 195) as a tense marker, though one, as Faraclas points out, "borrowed from the modality auxiliary system".) These markers are not obligatory. Where time reference is not indicated by preverbal markers or by time adverbials or other contextual cues, the verb is said to receive a non-past reading if it is stative, and a past reading if it is non-stative (Agheyisi 1971: 133f.; Faraclas 1996: 177). In addition to the preverbal markers for past, future/irrealis, and imperfective and perfective aspect, NigP has a preverbal negator which combines the meanings of negation and perfective aspect, *never* (cf. Agheyisi 1971: 148f.; Faraclas 1996: 89).[5] Faraclas (1987, 1996: 208f.) also reports an auxiliary use of *come*. Poplack and Tagliamonte (1996) confirm this analysis, but the

5. For an example, see text sample 3b, l. 5.

meaning of preverbal *come* is controversial. Faraclas (1987: 50; 1996: 208) interprets it as a marker of "realis modality" and explicitly asserts (1987: 49) that it indicates neither location in time nor temporal sequence, whereas Poplack and Tagliamonte describe it as a marker of "sequential past" (1996: 86). In the present corpus, a sequential analysis of preverbal *come* is possible in many cases, but the time reference of VPs with *come* is not always to the past. Rather, *come* can also combine with *go* to express sequentiality in non-past contexts (cf. e.g. text sample 1, ll. 15–17).

In addition to the NigP tense/aspect markers described above, the following English forms are attested in the corpus: the auxiliaries and regular inflections used in English for tense/aspect marking (*be, have, -ing, -ed*), irregular past tense/past participle forms,[6] and the past habitual marker *used to.* In Table 3, the number of verbs marked by each of these items, or combinations of markers where they occur, is compared to the number of verbs marked by NigP *dey, don, go, been* or *come* as well as combinations of these. As another option which distinguishes the NigP from the English system, unmarked verbs in past reference contexts are also taken into account. Furthermore, instances of the modal *will* (also *shall*) have been counted in order to enable a comparison with the NigP future marker.[7] Note, however, that while *will* often translates as preverbal *go* in NigP, English perfect forms as *don/never* + V etc., there is no one-to-one correspondence between any two categories in the table. The data thus do not represent results of variable analyses in the strict sense. As has been shown for example by Winford (1984), such analyses are problematic when the variant realizations of potential morphosyntactic variables are part of widely divergent systems like those of a pidgin/creole and its lexifier. This analysis (and also the one of copular verbs and constructions in 5.2) is therefore simply one of the extent to which non-NigP forms in a particular area of the grammar are used in the corpus in comparison to the set of NigP marking options, without claims of direct equivalence. Not included in these data are a number of cases in which the options for tense/aspect marking are restricted: VPs in imperative clauses, *make*-subjunctive clauses (cf. Faraclas 1996: 24ff., 209ff.), *to*-infinitive clauses and *-ing* clauses, non-initial members of serial constructions and VPs introduced by modal verbs (except for the combinations with *go* listed in the table). Non-NigP tense/aspect forms in these contexts are limited to a number of occurrences of V*ing* in nonfinite clauses, most of them in parts of the corpus marked as non-NigP speech and other English discourse chunks or idiomatic expressions in English that had been identified by the informants in the questionnaire study (cf. Section 4).

The first important finding to emerge from this table is that the data fall neatly into two categories, namely on the one hand forms which are fully grammatical according

6. Note that there are a few cases where NigP has adopted an English irregular past tense/past participle form as its invariant form (*lef* 'leave', *loss/lost* 'lose', *broke* 'break', *born* 'give birth to'). These forms therefore do not count as inflected forms.

7. In contrast to *will*, *be going to* does not occur in the corpus at all.

Table 3. Tense/aspect marking and modals go and will/shall in the NigP corpus

Forms	Number of occurrences*	
inflected *be* + V*ing*	12	[8] (1)
(inflected) *have* + past participle	1	[1]
used to + V	1	
V*ed* (regular inflected past)	8	[5]
irregular inflected past	7	[1]**
will/'ll/won't	13	[5] (4)
shall	4	[1]†
dey + V	1468	
don + V	648	
don dey + V	55	
don come + V	6	
never + V	113	
go + V	1520	
go dey + V	133	
go don + V	12	
go come + V	54	
go come dey + V	4	
been + V	7	
been dey + V	4	
come + V	233	
come dey + V	75	
zero past‡	1112	
TOTAL	5487	

* Numbers indicated in square brackets refer to tokens (out of the total number) in parts of the corpus marked as non-NigP speech and numbers indicated in round brackets to those in other discourse chunks or idiomatic expressions in English as cited in the questionnaire study (cf. Section 4).

** Of the six occurrences outside non-NigP speech, two are in self-correction contexts (cf. e.g. ([*?he?*] *brought*) *e bring di person from anoder place* 'he brought the person from somewhere else' (C10 112–113)).

† The three occurrences outside non-NigP speech are all accounted for by one sentence containing *shall* which is repeated three times (R06 212–214).

‡ Excluding regular verbs in phonetic environments which make it impossible to determine the presence or absence of the past tense morpheme, verbs with identical present and past tense forms, modal verbs and lexical items corresponding to English adjectives used as verbs.

to the standard English system and, on the other, the traditional NigP marking options. Unlike in the Jamaican mesolect or the Guyanese one (cf. Bickerton 1975b: 73ff.), there are no intermediate forms like those in the example given in Section 1, viz, the *-ing* suffix for progressive rather than a preverbal marker, but without auxiliary *be* or with the form *is* of the auxiliary for all persons; nor does NigP have a mesolectal replacement for preverbal *been* like *did* in Jamaican and Guyanese Creole (cf. Patrick 1999: Chapter 6; Bickerton 1975b: 69ff.). Secondly, it is noteworthy how frequent all

Table 4. Copulas and related constructions in the NigP corpus

Forms/constructions		Number of occurrences*	
Inflected *be* (copula):**	*is/'s*	121	[37] (41)
	are/'re	20	[7] (8)
	was	8	[3]†
	'm	1	
na (cleft sentence introducer and equative copula)		1648	
Uninflected *be* (equative copula)		870	
dey (locative copula and existential verb)		499	
Lexical items corresponding to English adjectives and quantifiers functioning as verbs (no copula)		246	
Impersonal *e get* (existential construction)		60	
TOTAL		3473	

* Numbers indicated in square brackets refer to tokens (out of the total number) in parts of the corpus marked as non-NigP speech and numbers indicated in round brackets to those in other discourse chunks or idiomatic expressions in English as cited in the questionnaire study (cf. Section 4)

** Inflected forms of copular *be* other than those listed do not occur.

† Of the five occurrences outside non-NigP speech, two are in a self-correction context: (*di result was- di result was- e dey*) *e dey okay sha//* (C06 21).

of the NigP options (except preverbal *been*) are in comparison to the English ones.[8] The very low degree of variation observed here may be contrasted for example with the results of Patrick's (1999) analysis of past marking in his speech data from predominantly mesolectal Jamaican Creole speakers: 34% of past reference verbs in his data have English inflectional past marking (see p. 209). Thus, according to Patrick (1999: 222), "what distinguishes past-reference in the JC [Jamaican Creole] mesolect from standard and non-standard varieties of English is first and foremost the existence and extent of variable marking". Lastly, not only are English forms very infrequent in the present corpus, but, as indicated in Table 3, a considerable proportion of those that do occur are part of discourse chunks or idiomatic expressions wholly in English as identified through the questionnaire study.

5.2 Copulas and related constructions

The results for this area of the grammar (Table 4) are similar to those for tense/aspect marking.

8. In Poplack and Tagliamonte's (1996: 97) data for past marking in educated NigP as spoken by Nigerian immigrants in Canada, preverbal *been* (spelled *bin* in Poplack and Tagliamonte's study) is also rather infrequent. For other varieties of NigP, quantitative data are not available in the literature.

While English inflected copula forms (see the upper part of the table) are more frequent than the English tense/aspect forms in Table 3, the proportion of occurrences in whole discourse chunks or idiomatic expressions in English cited in the questionnaires is again very high, and on the whole the forms and constructions which in NigP fill the semantic space of the copula *be* in English, listed in the lower part of the table, clearly predominate. Another important finding is that whenever non-NigP forms are used, they are, again, fully grammatical according to the English system, i.e. person/number concord is always observed. There is no form like generalized *is* without person/number concord, as in the mesolects of English-based Caribbean creoles (cf. Bickerton 1973: 651; Holm 1988: 176). Furthermore, one does not find zero before noun phrases and locative complements; copula use in these environments is obligatory. This contrasts with the situation in the Guyanese continuum as described by Bickerton (1973: 652ff.). According to Bickerton, zero copula in these environments is typical of the mesolect, while the basilect has the equative copula *a* and the locative copula *de*, and the acrolect in both cases inflected *be*.

5.3 Verbal negation

Verbal negation is also an area where special mesolectal forms occur in Caribbean creole continua. For example, *duon* in Jamaican Creole and *en* in Guyanese Creole can replace basilectal preverbal *no* and *na,* respectively, whether the time reference is past or non-past,[9] thus differing from forms like standard English *don't* and nonstandard English *ain't* or standard English *haven't* (Patrick 1999: 199; Bickerton 1975b: 99f.; Rickford 1983: 314).[10]

In the present corpus there are only 10 instances of *don't*, all in contexts where this would be the appropriate form of verbal negation in English as well. There are also a few other instances of English verbal negation, where English copula forms or auxiliaries are negated by *not*, but such cases are not only infrequent but also largely confined to whole discourse chunks or idiomatic expressions in English cited in the questionnaires. In the overwhelming majority of cases the NigP pattern of verbal negation is applied, i.e. the verb is negated by *never* if it is also to be marked for perfective aspect (cf. 5.1), otherwise by preverbal *no*, of which there are over 1500 instances in the corpus.

---

9. According to Bickerton (1975b: 100f.), *doon* is also used in the Guyanese mesolect with both present and past reference. Rickford (1983: 314), however, argues that if forms like *don*(*t*), *doesn*(*t*) and *didn*(*t*) occur in Guyanese Creole, they almost always have the time reference associated with the corresponding English forms.

10. Bickerton (1975b: 91) says about the etymology of *en* in Guyanese Creole that "it is unclear whether *en* stems from *ain't, or haven't,* or is the joint product of both".

6. Interpretation of results

The quantitative results provided in the preceding section have shown that traditional NigP verbal structures, i.e. those that would be described as basilectal in a continuum, predominate even in the speech of the most educated users of the language. The corpus analysed did not provide any evidence of mesolectal forms of the type that occur in Caribbean creole continua. English verb forms, the acrolectal ones in a continuum, were found to be present in the corpus but it became clear that their use is very limited in comparison to that of the NigP forms and that they tend to occur mainly in special contexts where they are part of larger English segments.

If we go back to the text samples from the corpus in Section 4, we can see the results summarized above reflected there. In text sample 1, NigP verb forms and constructions are used throughout in spite of the heavy English lexical influence; cf. e.g. the frequently occurring tense/aspect markers *go* and *dey*, the two instances of *go come* already drawn attention to in 5.1, the unmarked past tense form *surprise* (l. 20), the copula in *dis thing* <u>*na*</u> *mathematics* (ll. 6–7) or *dis* <u>*na*</u> *di topic* (l. 10), and the two instances of verbal negation by *no* in ll. 31–32. It is also interesting that, as pointed out in Section 4, the pattern of English lexical influence observed here parallels Agheyisi's "language interlarding" or "the special kind of mixed code which results from the interspersing of indigenous language speech with English elements" (Agheyisi 1977: 97), a phenomenon which "has virtually become a stable part of the verbal repertoire of urban communities in Nigeria" (ibid.). In text samples 3a–d, the NigP tense/aspect markers, copula forms and negators described in Section 5 are also employed consistently except in the English segments that have been interpreted in Section 4 as the result of code-switching as a discourse strategy, where fully grammatical English forms are used – cf. the negated copula in 3c, ll. 4–5 and the past and past progressive forms in 3d, ll. 11–13. Code-switching into English as a discourse strategy is also widely attested in the indigenous language use of educated Nigerians (cf. e.g. Amuda 1994; Nwoye 1993 and Madaki 1983 for examples involving each of the three major indigenous languages).

The data thus suggest that NigP-English contact phenomena in the area of verbal structures have more in common with language contact phenomena in the English-influenced urban or educated varieties of indigenous languages of Nigeria than with the mesolectal structures characteristic of Caribbean creole continua.

7. Conclusion

It has become clear that from the point of view of the verbal structures examined in this paper, NigP and English can be described as two separate systems (like Yoruba and English, for instance), which alternate to some extent in the language use of bilinguals. They do not form opposite poles of a continuum of linguistic variation as found in the anglophone Caribbean. Why should this be so more than 40 years after Nigeria's in-

dependence, which, according to the theory mentioned in Section 2, created the kind of social conditions conducive to a significant change in this situation? There are different factors to be considered. Firstly, it may be significant that to the present day the contact situation in which NigP is involved has included not only its lexifier but also its substrates (cf. also Agheyisi 1984: 230f.). With these, NigP shares important structural features that distinguish it from English, not least in the verb phrase (cf. Faraclas 1988). Secondly, the theory that continua like the Jamaican or Guyanese one developed out of an earlier state with only two separate varieties and that the same development is bound to take place elsewhere may itself be flawed. In fact, in recent times, alternative explanations have been gaining ground. According to these, the mesolects that link the basilectal creole and the acrolect in creole continua go back right to the early days of contact between speakers of the substrates and speakers of the lexifier and the development of mesolectal varieties is in some way connected with the fact that some substrate speakers had relatively closer contacts with speakers of the lexifier than others. Explanations of this type have been put forward in different forms by creolists with such diverse orientations as Alleyne (1971, 1980), Mufwene (e.g. 2001), and also Bickerton himself, who, in the early 1980s (cf. Bickerton 1983), was converted to the position that instead of the mesolect resulting from decreolization of the basilect,

> the creole continuum came into existence 'backwards', so to speak – those varieties closest to English originating from the earliest contact, and those furthest from English, from the phase in which the original model was most drastically diluted by a massive and rapid increase in the non-European population.
>
> (Bickerton 1988: 272)

If some such process was involved in the development of creoles in the former settlement colonies (Mufwene 2001) of the Caribbean, it is not obvious that the linguistic outcome of the contact between substrate and superstrate speakers should have been the same in the former trade colonies (ibid.) of West Africa and Melanesia. Though expanded pidgins like NigP and Tok Pisin function like creoles today, they developed under very different conditions:

> Settlement colonies started with intimate interactions between the two parties. Segregation was subsequent to the increase in the sizes of the European population and the larger proportions of non-Europeans. Multilingualism led the Africans to adopt the languages of the groups in power as their vernaculars. These were restructured during the appropriation process. Trade colonies were characterized by random contacts between the European traders and their African counterparts. The adoption of European languages under these conditions of limited, occasional exposure to them allowed the development of what have also been identified as 'broken languages', a reflection of the minimal uses to which elements of their lexifiers were put. (Mufwene 2001: 171)

The same issue has been raised by Siegel in a paper entitled "Pidgin and English in Melanesia: is there a continuum?" (1997), a question which he answers in the negative.[11] He argues:

> The absence of extensive depidginization or decreolization in Melanesia may also support the view that a post-creole continuum is not the result of decreolization, but reflects a spectrum of variation that has existed since much earlier stages of development (Alleyne 1971; LePage 1977). There is no historical evidence of such variation in Melanesia (although it may have existed in Australia, where Melanesian Pidgin began to stabilize on the sugarcane plantations, and where contact between English-speakers and indentured Melanesians was more extensive). Thus, according to this view, we would not expect to find a pidgin-to-English continuum in Melanesia because there never has been one. (Siegel 1997: 201)

While the full complexity of the sociohistory of West African Pidgin cannot be unravelled here (but cf. Huber 1999), one could argue that by the same token, one would not expect a NigP-to-English continuum, whose existence has so far been both denied and asserted. The present paper is too limited in its scope to resolve the issue definitively, but the aspects of variation in educated NigP that have been studied here certainly support this view.

References

Agheyisi, R. N. (1971). West African Pidgin English: Simplification and simplicity. Ph.D dissertation, Stanford University. Ann Arbor MI: UMI.

Agheyisi, R. N. (1977). Language interlarding in the speech of Nigerians. In P. F. Kotey & H. Der-Houssikian (Eds.), *Language and Linguistic Problems in Africa* (pp. 97–110). Columbia: Hornbeam.

Agheyisi, R. N. (1984). Linguistic implications of the changing role of Nigerian Pidgin English. *English World-Wide, 5*, 211–233.

Alleyne, M. C. (1971). Acculturation and the cultural matrix of creolization. In D. Hymes (Ed.), *Pidginization and Creolization of Languages* (pp. 169–186). Cambridge: CUP.

Alleyne, M. C. (1980). *Comparative Afro-American: An historical-comparative study of English-based Afro-American dialects in the new world.* Ann Arbor MI: Karoma.

Amuda, A. A. (1994). Yoruba/English conversational code-switching as a conversational strategy. *African Languages and Cultures, 7*, 121–131.

Bickerton, D. (1973). The nature of a creole continuum. *Language, 49*, 640–669.

Bickerton, D. (1975a). Can English and Pidgin be kept apart? In K. A. McElhanon (Ed.), *Tok Pisin I Go We?* (pp. 21–27). Port Moresby: Linguistic Society of Papua New Guinea.

Bickerton, D. (1975b). *Dynamics of a Creole System.* Cambridge: CUP.

11. Cf. in this connection also Smith (2002), specifically on the case of Papua New Guinea. Smith concludes from his data analyses that "the current situation cannot be characterized as a continuum", as "there is definite discontinuity between Tok Pisin and English in the huge majority of cases" (p. 210).

Bickerton, D. (1983). Review of Baker and Corne, *Isle de France Creole*. *The Carrier Pidgin, 13*(4), 8–9.

Bickerton, D. (1988). Creole languages and the bioprogramme. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge survey,* Vol. 2: *Linguistic Theory: Extensions and implications* (pp. 268–284). Cambridge: CUP.

DeCamp, D. (1971). Toward a generative analysis of a post-creole speech continuum. In D. Hymes (Ed.), *Pidginization and Creolization of Languages* (pp. 349–370). Cambridge: CUP.

Deuber, D. (2002). 'First year of nation's return to government of make you talk your own make I talk my own': Anglicisms versus pidginization in news translations into Nigerian Pidgin. *English World-Wide, 23*, 195–222.

Deuber, D. & Oloko, P. (2003). Linguistic and literary development of Nigerian Pidgin: The contribution of radio drama. In C. Mair (Ed.), *The Politics of English as a World Language: New horizons in postcolonial cultural studies* (pp. 289–303). Amsterdam: Rodopi.

Elugbe, B. (1995). Nigerian Pidgin: Problems and prospects. In A. Bamgbose, A. Banjo, & A. Thomas (Eds.), *New Englishes: A West African perspective* (pp. 284–306). Ibadan: Mosuro.

Faraclas, N. (1986). Pronouns, creolization and decreolization in Nigerian Pidgin: A pilot study. *Journal of West African Languages, 16*(2), 3–8.

Faraclas, N. (1987). Creolization and the tense-aspect-modality system of Nigerian Pidgin. *Journal of African Languages and Linguistics, 9*, 45–59.

Faraclas, N. (1988). Nigerian Pidgin and the languages of southern Nigeria. *Journal of Pidgin and Creole Languages, 3*, 177–197.

Faraclas, N. (1996). *Nigerian Pidgin.* London: Routledge.

Grimes, B. (Ed.). (2000). *Ethnologue: Languages of the world.* (14th edition). Web edition: http://www.ethnologue.com. Dallas: SIL.

Gumperz, J. J. (1982). *Discourse Strategies.* Cambridge: CUP.

Holm, J. A. (1988). *Pidgins and Creoles,* Vol. 1: *Theory and structure.* Cambridge: CUP.

Huber, M. (1999). *Ghanaian Pidgin English in Its West African Context: A sociohistorical and structural analysis.* Amsterdam: John Benjamins.

Madaki, R. O. (1983). A Linguistic and Pragmatic Analysis of Hausa-English Code-Switching. Ph.D dissertation, University of Michigan. Ann Arbor MI: UMI.

Mann, C. C. (1993). The sociolinguistic status of Anglo-Nigerian Pidgin: An overview. *International Journal of the Sociology of Language, 100–101*, 167–178.

Mann, C. C. (1998). Language, mass communication and national development: The role, perceptions and potential of Anglo-Nigerian Pidgin in the Nigerian mass media. In *Language in Development: Access, empowerment, opportunity* (pp. 136–144). Kuala Lumpur: National Institute of Public Administration.

Mufwene, S. S. (2001). *The Ecology of Language Evolution.* Cambridge: CUP.

Nwoye, O. G. (1993). Code-switching as a conscious discourse strategy: Evidence from Igbo. *Multilingua, 12*, 365–385.

Patrick, P. L. (1999). *Urban Jamaican Creole: Variation in the mesolect.* Amsterdam: John Benjamins.

Poplack, S. & Tagliamonte, S. (1996). Nothing in context: Variation, grammaticization and past time marking in Nigerian Pidgin English. In P. Baker & A. Syea (Eds.), *Changing Meanings, Changing Functions: Papers relating to grammaticalization in contact languages* (pp. 71–94). London: University of Westminster Press.

Rickford, J. R. (1983). What happens in decreolization. In R. W. Anderson (Ed.), *Pidginization and Creolization as Language Acquisition* (pp. 298–319). Rowley MA: Newbury House.

Sebba, M. (1997). *Contact Languages: Pidgins and creoles.* Basingstoke: Macmillan.

Shnukal, A. & Marchese, L. (1983). Creolization of Nigerian Pidgin English: A progress report. *English World-Wide, 4*, 17–26.

Siegel, J. (1997). Pidgin and English in Melanesia: Is there a continuum? *World Englishes, 16*, 185–204.

Smith, G. P. (2002). *Growing up with Tok Pisin: Contact, creolization and change in Papua New Guinea's national language.* London: Battlebridge.

Tagliamonte, S., Poplack, S., & Eze, E. (1997). Plural marking patterns in Nigerian Pidgin English. *Journal of Pidgin and Creole Languages, 12*, 103–129.

Todd, L. (1974). *Pidgins and Creoles*. London: Routledge.

Winford, D. (1984). The linguistic variable and syntactic variation in creole continua. *Lingua, 62*, 267–88.

A linguistic time-capsule

Plural /s/ reduction in Afro-Portuguese and Afro-Hispanic historical texts

Fernanda L. Ferreira
Bridgewater State College, USA

The present study builds on the work of Lipski (2005a) regarding the nature and scope of Afro-Iberian language. It narrowly focuses on pluralization patterns found in historical texts covering five centuries. The analyses show that the representations of Black speech in Afro-Portuguese examples are consistent with a pattern of plural marking independently verified in comparable creole and semi-creole systems, such as Cape Verdean, Papiamentu, Helvécia and popular varieties of Brazilian Portuguese. Comparatively, the evidence of variation found in earlier Afro-Hispanic texts is less conclusive about the possible contribution of the *bozal* language to modern-day varieties of Caribbean Spanish. The analyses add to the debate over the prior creolization of non-standard varieties in Latin America.

1. Introduction

In his book *A History of Afro-Hispanic Language: five centuries/five continents*, Lipski (2005a) discusses in great detail the nature and scope of Afro-Iberian language spoken by Blacks starting from the 15th to the 20th century. He does so using historical textual data. The present article narrowly focuses on pluralization patterns found in the texts originally compiled by that author.[1]

The dating of phonological changes before the advent of recording technology has been a methodological problem for dialectologists and historical linguists. One of

1. In writing this article I am indebted to the groundbreaking work of John M. Lipski, who taught me and many others about the world of Afro-Hispanic language in several riveting class sessions at the University of New Mexico. A paper version of this article was presented at the SPCL annual conference in January 2003 with support from Bridgewater State College's Center for the Advancement of Research and Teaching (CART). All remaining shortcomings are, of course, mine.

the methods employed to circumvent this problem is the analysis of historical texts in order to shed light on phonological as well as morphological and syntactic changes in language. There are other methods also employed to study the development languages, namely, the study of related varieties (i.e. Sephardic Spanish). However, in the absence of conservative linguistic systems that present evidence of particular consonantal changes, researchers need more examples of what one may call 'a linguistic time capsule'. Just as with time capsules left by previous generations to inform future ones of the details of the past, written texts may provide partial evidence of the intricacies of language left by authors of the past.

With respect to phonetic changes, there are obvious limitations to the written text. In languages such as Chinese, for example, the use of ideograms provides no phonetic information at all (Labov 1994: 11). Even in languages where these limitations do not occur, there are still obstacles. Subtle allophonic detail is obscured in textual representation, such as the difference between Portuguese mid vowels [e] and [ɛ] both graphically represented by <e>[2] Nevertheless, in the case of Spanish and Portuguese, researchers are in a fortunate position with regard to a number of phonological features: <r> and <l> interchanges (*farta* < *falta* 'lack of') and epenthetic processes (*dioso* < *Dios* 'God') are generally represented in the written text.

Still, the representation of speech in literary sources remains problematic. Contemporary portrayals of Black speech in film, for example, have been reported to be quantitatively, and to a lesser extent qualitatively, different from variation patterns in naturally occurring speech (Wilkerson 2000).[3] In the case of the present study, literary sources are used to uncover orthographic clues that can be used to establish a hypothetical path of development, keeping in mind that the technique may depend on the sociolinguistic sensibility of the writer, i.e. the author's attentiveness to the way in which people spoke and accuracy in its depiction.

Possible scribal error is also mentioned as another limitation in the analysis of written texts. However, with the advent of the printing press, scribal error ceases to be an obstacle (Lipski, p.c.). In addition, the question of the writers' backgrounds is important, since written materials of the past were usually a product of the best efforts of a learned few. The question of perceived errors in language and subsequent language change is important, since these changes can be classified as 'change from below' (Labov 1994: 78–79), in which linguistic processes take place first among speakers with less formal education and from lower social strata. It is also possible to have 'change from above', in which innovative processes start by the more educated layers of society. It is improbable that grammatical patterns considered incorrect would be present in a literary text, since it would have been produced by educated people. Nevertheless, lit-

2. For example *beco* [be.ku] 'alley' and *queto* [kɛ.tu] 'quiet, calm', although diacritcs can help distinguish these vowels sometimes (cf. *café* [ka.fɛ] 'coffee' and demonstrative *este* 'this' [es.ti]).

3. I am thankful to the anonymous reviewer of this paper, who alerted me to the work of Rose Wilkerson in African American English and film.

erary sources, especially plays where dialogues of speakers from marginalized groups are incorporated, can be a useful tool to examine paths of development in a language. Thus, linguistic research into historical texts, as stated by Labov (1994: 11), is "the art of making the best use of bad data." In this study, I show that the data selected are helpful in uncovering grammatical patterns and variations of plural /s/ reduction in Afro-Iberian language. Clearly, using historical texts is not innovative in and of itself, but in this particular case, it adds to the overall picture of phonological variation in dialects of Spanish and Portuguese.

2. The first act: 15th and 16th centuries

In the next sections I present evidence of plural /s/ deletion found in Afro-Portuguese and Afro-Hispanic texts starting from the 15th up to the early 20th century. The beginning stage refers to the early examples of Portuguese and Spanish literature. Table 1 lists the authors and titles of texts from the 15th and 16th century included in this study.

As with many early authors of the 16th century, the birth date of Fernam (sometimes spelled *Fernão*) da Silveira is not exactly known. He had been the poet in the court of João II, and later fell from grace and became an enemy of the king. He escaped from the court and sought refuge first in Spain and later in France, where he was killed in 1489 (Esteves Pereira 1904–1915). Despite his misfortunes, he remains one of the pioneers of the Portuguese letters.

Gil Vicente was born between 1460 and 1470, the specific date, again, still a matter of speculation. A Portuguese poet and dramatist, he served the courts of the Portuguese kings Manuel I and João II. Garay (1988) adds that Gil Vicente's plays "vividly

Table 1. 15th and 16th century texts included in the present study

Afro-Portuguese	Afro-Hispanic
– Fernam da Silveira (?–1489) *A min rrey de negro*[...] (Guimarais 1909–1917)	– Rodrigo de Reinosa (?1480–?1520) *Copla* (Cossío 1950)
– Gil Vicente (?1470–?1536) *Clérigo da Beyra* *Nao d'Amores* (Vicente 1834, 1912)	– Sánchez de Badajoz (?1490–?1552) *Farsa teologal* *Farsa del moysen* (Barrantes 1886)
– António Ribeiro Chiado (?1520–1591) *Prática de Oito Figuras* (Chiado 1961; Pimentel 1901)	– Lope de Rueda (1510–1565) *Comedia de los Engañados* *Comedia de Tymbria* *Comedia de Eufemia* (Rueda 1908)
– Sebastião Pires (?1480–?) *Auto da bella menina* (Pires 1922)	

portray the breadth of Portuguese society". It is not understood exactly why Vicente produced literary texts in both Spanish and Portuguese, but it is clear that he was important in shaping both modern Spanish and Portuguese drama.

Born in Évora, António Ribeiro Chiado (?1520–1591), often called simply Chiado (a topographical reference), was a well-known writer of verses in the 16th century. His works include *autos* (short plays) and religious plays that followed the style of Gil Vicente. He was an expert at improvisations and was able to perfectly imitate the voice and gestures of famous figures of the time (Esteves Pereira 1904–1915). His comedic artistry made him a popular artist in Lisbon during his lifetime.

Very little is known about Sebastião Pires (?1480–?), other than he contributed *autos* that are very similar in style to those included in the collection of courtly poetry named *Cancioneiro Geral Garcia de Resende* published in 1516 (Azevedo 2005; Lipski 2005a).

Similarly to authors of the time, there is very little concrete biographical information on Rodrigo de Reinosa (?1480–?1520). His name was probably topographical, meaning, Rodrigo, 'the one born in (the area of) Reinosa.'

Diego Sánchez de Badajoz (?1490–?1552) was the author of various *autos* and religious *farsas* that included some picaresque or satirical elements. After his death, his nephew published a compilation of his theatrical works that appeared in 1554. His plays are moralistic and religious in nature, interspersed with comic relief by the introduction of well-known characters such as 'the Black', 'the Priest', or 'the Soldier' (Marbán 1971).

Lope de Rueda (1510–1565) was born in Seville, and lived in Valladolid, Spain. While in that city he constructed a temporary playhouse where he had his plays performed. As an actor and director of his own plays, he introduced innovations to the theater, and his theater company traveled all over Spain. His works had a popular style that appealed to the public, and he became famous for his *pazos*, short plays used to entertain the audience before a longer play. Again, the protagonists in these short plays were caricatures of everyday Spanish life, such as 'the Servant', 'the Bachelor', and so on. He is considered to be the founder of Spanish theater (Buendia 1965).

In the next two sections I will discuss first the Afro-Portuguese texts and then the Afro-Hispanic texts listed in Table 1.

3. Early Afro-Portuguese texts (15th and 16th centuries)

The authors listed above wrote representations of language used by Black speakers, known as *língua de preto* (in Portugal) or *habla de negro* (in Spain). These authors attempted to represent the speech of Blacks living in the southern region of the Iberian

Peninsula, usually in the cities of Lisbon and Seville.[4] With the importation of Black Africans, the Iberian peninsula acquired a population of Africans who were forcibly brought to Europe as slaves, domestic servants, or more generally, forced-labor workers. They later acquired freedom and some formed trade societies, *cofradías* or brotherhoods. They were forced to learn either Portuguese or Spanish under considerably unfavorable conditions. Moreover, these Blacks, who spoke either Portuguese or Spanish as a second language, were referred to as *bozales*, from the word in Spanish roughly meaning 'untamed, savage' (Lipski 2005a). As a consequence of the presence of these African populations in the Iberian peninsula, the character of the Black African appears in the early literature of both Portugal and Spain (Lipski 1998). Also according to Lipski (1995: 133) the earliest attestation of *bozal* language is present in the 15th century texts of the collection known as *Cancioneiro Geral Garcia de Resende*, and later in the writings of Gil Vicente (1470?–1536?) and António Ribeiro Chiado (?–1591). Example (1) is from Fernam da Silveira (?–1489) found in the *Cancioneiro* while examples (2) through (5) are from Vicente and Chiado.

(1) Excerpt of *A min rrey de negro* [..], ca. 1455. (Lipski 2005b: 2; Guimarais 1909–1917)

A min rrey de negro estar Serra Lyoa	I am the king of Blacks of Sierra Leone
lonje muyto terra onde viver nos,	We live in a land far away
andar carabela, tubao de Lixboa,	Go/travel by ship, [?] from Lisbon
falar **muyto novas** casar pera vos.	Speak **many news** [?] for you
Querer a mym logo ver-vos como vay;	I soon want to see/know how you go
Leyxar molher meu, partir, muyto synha,	leave my wife, go away, very [?]
Porque sempre nos servyr vosso pay,	because your father has always served us
Folgar **muyto negro** estar vos rraynha.	freed **many Blacks** your queen has

The first noun phrase in bold in example (1) is somewhat ambiguous since the word *muyto* (cf. *muito* in standard Portuguese) could be translated as either the modifier 'much' usually followed by an adjective (cf. *muito novo* 'very new/young') or as the adjective/modifier 'many' followed by a noun (cf. *muitas novas* 'many news'). If one assumes that the head of the phrase is the word *novas*, a plural noun, the gloss 'many news' can be accepted. By contrast, if one assumes that the head of the phrase is the adjective 'young' (*nova*) then it can be glossed as 'very young'. But the rest of the poem includes other examples of *muyto*, which are followed by, according to the present interpretation, nouns such as *synha* and *negro*. For those reasons, the interpretation of *novas* as an adjective rather than a noun is not warranted. It should be noted that

4. However, concentrations of Blacks were also found as far north as Madrid, Valencia, as well as in Galicia, Extremadura and the Basque country (Lipski 1998: 301, my translation).

the previous verb *falar* 'to speak' seems to indicate that the best gloss for the noun phrase is indeed 'many news', as in 'to speak or convey many news'. In his analysis, Lipski (2005: 59) notes that "the poem provides limited opportunities for noun-adjective concord, but several lapses are noted". The last example *muyto negro* is interesting because even though there are no overt plural markers in the surface form, this is the case of a plural generic referent, which is standard use in current Portuguese. Its significance rests on fact that when counting plural forms, researchers have to account for the various representations of plural notions, such as in this one.

The poem as a whole describes the voyages of Vasco da Gama to the African continent, and this particular stanza refers to the speech of a tribal king from Sierra Leone. In it, there are examples of the use of infinitives instead of inflected verb forms (*onde viver nos* 'where we live') and the form *mim* (standard Portuguese *mim* 'me' is a prepositional object) used as a subject pronoun (*A min rrey de negro estar Serra Lyoa* 'I am King of Blacks of Sierra Leone'). These features also occur in present-day Afro-Portuguese creoles (Mello 1996). There are different plural marking patterns found in popular or non-standard varieties of Romance languages (Holm 1988–89) which differ from the above text. In these varieties, plural /s/ is marked on the first element of the noun phrase, usually a determiner, and is absent at the end of the noun (a 's_ø' pattern). In the case of the text above, this particular pattern is not followed. In fact, the marker is present at the end of the noun, but absent at the end of the determiner (a 'ø_s' pattern).

Examples (2) and (3) are from Gil Vicente (Lipski 2005b: 3–8), who was responsible for the largest share of the Afro-Portuguese literary corpus, with plays written in the 1520's and 1530's. As mentioned earlier, this author is particularly important because he wrote in Portuguese and in Spanish, establishing a connection between Afro-Portuguese and Afro-Hispanic literary traditions. Again, the reasons for which Vicente produced an almost equal number of plays in Portuguese and Spanish is still unknown.

(2) Excerpt of *O Clérigo da Beyra* (Lipski 2005b: 3; Vicente 1907)

Boso seria muito bó;	You would be very [good?]
vaca ne Francico paia;	a cow [not?] Francisco [?]
tenha **seis filho** e mi so	to have **six sons** and I only/barely have
nam temo comere ni migaia...	not one morsel to eat...

(3) Excerpt of *Nao d' Amores* by Gil Vicente (Lipksi 2005b: 8; Vicente 1834)

A da elle roma	To give s/he [?]
doze, que a mi comprae.	Twelve, that I buy
e masa **cinco maçao**	and more? **five apples**
Se a mi vae elle falae	If I go to talk to her
faze carneo de verao	to make [?] of the summer

In examples (2) and (3) the noun phrases are composed of a numeral adjective followed by a noun lacking plural marking. There is an apparent logic to this pattern since

the semantic load of plurality is already placed on the numeral element, and markers are not necessary in subsequent elements to convey the idea of plurality. Indeed, Seklaoui (1986: 51), in discussing parallel weakening of /s/ in French, Italian and Spanish, mentions that the /s/ morpheme is retained more in the initial position of the noun phrase, because it would be the first element to transmit the information of plurality. She adds: "following this hypothesis, nouns would not need to preserve the redundant information of the plural morpheme".[5] Kiparksy (1982) had put forth the notion of "distinctiveness condition" which assumes that morphemes carrying relevant semantic information are less likely to be deleted, following a functionalist framework. In that view, one could argue that the /s/ at the end of nouns carries relevant information and is thus less likely to be deleted. On the other hand, one could argue that if the semantic information has already been conveyed (by the first element in the noun phrase) then the marking on the noun itself is redundant rather than relevant. In his discussion of functionalism and the mechanisms of linguistic change, Labov (1994) explains how patterns tend to be repeated in discourse, regardless of a functional basis. In the case of the noun *maçao*, the author follows a less redundant pattern, which is repeated throughout the text.

Example (4) is a lesser-known text from António Ribeiro Chiado, circa 1550, reproduced here.

(4) Excerpt of *Prática de oito figuras,* (Lipski 2005b: 10; Pimentel 1889)

Doso gallia, um capao;	**two chickens**, one [?]
a mim traze turo junto;	I bring all together
o coeio, co' **treze pombio**...	the rabbit, **thirteen pigeons**
[...]	[...]
Nunca elle mem acha	Never does he find/think
muito caro, nunca bem,	very expensive, never good [?]
mim da-le **treze vintem**	I give him **thirteen *vintens*** [currency]
pr'o dizo nun quere dá	for [?] does not want to give

The two stanzas selected here show other creole-like features, such as the epenthetic vowel at the end of the word *dos* 'two', which follows a preference for the CV syllable. This syllable structure is common in Spanish and Portuguese-based creoles (Schwegler 1998: 261) and in the African substrate languages (Castro 1980: 18–19).[6] All the noun phrases included in the text are formed by a numeral adjective followed by a noun. Again, there are no plural /s/ markers at the end of nouns.

In the text entitled *Auto da bella menina* by Sebastião Pires (?1480–?), who lived approximately at the same time as Gil Vicente (Lipski 2005a: 58), there is considerably

5. This notion is also mentioned in Ma and Herasimchuk 1971 for Puerto Rican Spanish.

6. More on the issue of the syllable structure of Afro-Hispanic language and phonotactic adjustments can be found in Lipski (2005a).

more plural /s/ marking than in the texts of his contemporaries. Example (5) illustrates the increased number of plural marking.

(5) Excerpt of *Auto da bella menina,* (Lipski 2005b: 11; Pires 1922)

ou **la gentes** ou falay corpo na sam quebray	or **the people** speak the body [?] without breaking
home **sua dentes** o reca do **sua parentes** ou hora beiyo mao…	men and **their teeth** or [?] of **their relatives**, the time I kiss hand
[…]	[…]
y mias elle manda beyjar **suas dedos** co	and she tells me to kiss your fingers with the
coracanbar dessa caroza galante…	[?] in this gallant [?]…
negro nam sa gete e boso zamboy de mi.	the Blacks are not [considered] people and you make fun of me.
Eu **suas comendas**	I give [?] y**our orders** [?]
day que dele manda traze ca e com sua yrmao falay	that he tells (me) to bring and speak with …your brother

The lack of gender agreement in the noun phrase seems to be in concordance with features previously expressed in other Afro-Portuguese texts, such as paragogic vowels (*boso* < *vos*, 'second person singular pronoun'), use of *mi*(*m*) as subject pronoun, and use of infinitives instead of finite forms (*beyjar*). In addition, gender marking in these texts parallel those found in *Helvécia*, an isolated Afro-Brazilian community in Bahia (Baxter, Lucchesi & Guimarães 1997). However, plural /s/ is consistently found at the end of the noun phrase. In the last noun phrases of example (5) one finds nominal plural concord (cf. *suas comendas* and *suas dedos*), which should be noted. These are the only examples where such concord is found in the text of Sebastião Pires.

4. Afro-Hispanic texts (15th and 16th centuries)

This section deals with the works of Spanish authors of the same period as the previously mentioned Portuguese writers. They are Rodrigo de Reinosa, Sánchez de Badajoz and Lope de Rueda (see Table 1).

In Golden Age Spanish literature many common stereotypes, such as the *el moro* 'the Moor' and the *el gitano* 'the Gypsy', were attributed linguistic characteristics (Lipski 1995: 135). Comparable caricatures were done for the *negros bozales* of that time. Again, the *bozales* were Blacks born in Africa who spoke Spanish and Portuguese as a second language and whose speech is represented in various poems and plays. How accurate can one expect these representations to be? The answer depends on who wrote the text and how familiar he was with the speech of *bozales*, as well as how consistent he was with respect to those features.

The early representation of Black speech in Spain come from Rodrigo de Reinosa. His contact with a possible early Afro-Iberian pidgin spoken in the peninsula is important in analyzing his writings. He most probably came into contact with marginalized groups in Seville (Lipski 2005a). Although Reinosa could not have read attestations of Afro-Portugese language in the *Cancioneiro Geral* (published decades later) he might have been inspired by early popular pamphlets containing some Afro-Hispanic language (Lipski 2005a: 71–72). Therefore, Reinosa's writings are probably not imitation, but rather, inspiration from first-hand experience. The intended audience of Reinosa's 'africanized' *coplas* (verses of popular songs), according to Lipski, was a white working class, although it is possible that Blacks were also among them. Example (6) is from a *copla* by Rodrigo de Reinosa (?1480?–?1520).

(6) Excerpt of *Copla* (Lipski 2005b: 32; Cossío 1950)

A mí saber bien bailar el guineo	I know well how to dance the Guineo
si querer, conmigo facer choque, choque	if you to do with me the [dance?]
y con un bezul **dos veces** arreo	and with the [?] **two times** [?]
O me iba allá a la horta	I go there to the orchard
para el nabo mercar,	to get the turnip to sell?
e abrir pasico la porta	to open to [?] the door
y hortelano no fallar;	and not fail/ miss the mint
y **los nabos** le hurtar	and to steal **the turnips**

As can be surmised from the words in bold in example (6), there is no evidence of plural /s/ deletion in these literary representations of African speech. Either Reinosa was not particularly accurate in expressing a feature attested by his contemporaries or, in fact, the *bozales* he came in contact with in Andalusia did not exhibit lack of number agreement in their speech. To be sure, it is possible that contact varieties spoken by Blacks in the Iberian Peninsula at the time did not share many features. Another explanation rests on the possible confusion with Andalusian speech. Weakening processes for sibilants in coda position were already underway all dialects of Spanish (Seklaoui 1986), so one might assume that this particular trait could not be used in the literary genre to be representative of Black speech. It is important to note that other common features of Africanized varieties are present in these verses: the use of *mí* as subject pronoun and the use of verbal infinitives instead of the finite forms (*fallar*, *saber*) in the verb phrases. Thus, this author was accurate and consistent with regard to other linguistic phenomena reported to be characteristic of Africanized varieties at the time.

The next examples of literary texts come from Sánchez de Badajoz (?1490–?1552). This author wrote many *farsas* or short plays which are dated around 1533 and 1548 (Lipski 2005a: 73). Badajoz did not live near the southern cities of Spain, where the African-born speakers were mostly located. Nevertheless, he "demonstrated a keen awareness of popular speech" (Lipski 2005a: 74). Because of Badajoz's contact with rustic varieties, it is important to analyze the features presented here keeping in mind that, at the same time, they may be reproductions of rural speech (cf. discussion of

non-standard varieties particularly *sayagués* in Lipski 1995). There is some controversy regarding the authorship of the verses in examples (7) and (8), but the majority of scholars accept that they were written by Sánchez de Badajoz in the early to mid 16th century.

(7) Excerpt of *Farsa teologal*, (Lipski 2005b: 35; Barrantes 1886)

Mirara tam bonita...	Looked so beautiful
Varara **ras enemigos**,	[?] **by the enemies**
aunque das, bona fe ygos	Although in good faith [?]

(8) Excerpt of *Farsa del moysen*, (Lipski 2005b: 36; Barrantes 1886)

Lesa caye barados	Let fall [?]
ba por **dus ano** quirados,	Go for **two years** [?]
para bender econdido...	To sell unnoticed
Sura san de agua no hato	His [?] of water in the [?]

In example (7) the phrase *ras enemigos* is assumed to have derived from both the interchange of *r* and *l*, *ras* < *las*, and also the lack of gender concord, where *las* 'DET-plural-fem' *enemigos* 'enemies' is used where *los* 'DET-plural-masc' would be expected.

In example (8) the word *dus* is glossed as the monomorphemic 'two' (cf. *dos* in Spanish). Although monomorphemic words ending in /s/ were not specifically analyzed in this study, the presence of such a word in the noun phrase is important since phonological processes of /s/ deletion can be at work, rather than morphological ones. In the context of this noun phrase, it is important to note that the third element receives an overt plural marker. Lipski (2005a) argues that the difference between this author's texts and the Portuguese corollaries may be explained by the fact that the locally formed Afro-Hispanic language was gradually replacing the possible Afro-Portuguese pidgin that existed in Spain. In addition, the collection of *farsas* by Sanchez de Badajoz presents the first apparent examples of loss of syllable-final /s/ in Andalusia (cf. *etar* < *estar* 'to be', *apueta* < *apuesta* 'bet' as cited in Lipski 2005a: 75).

The last author to be analyzed here is Lope de Rueda (1510–1565). It can be said that this author contributed in large part to the peak of 16th century *habla de negro*. According to Lipski (2005a) the plays written between 1538 and 1545 are the most reliable and consistent examples of *habla de negro* found in 16th century literature. The quality and consistency of linguistic features found in the plays may be due to the fact that Lope de Rueda was in direct contact with African-born speakers of Spanish and may in fact have played some of the roles in the *comedias*. Examples (9) through (11) illustrate such texts.

(9) Excerpt of *Comedia de los Engañados* (Lipski 2005b: 38; Rueda 1908)

¿Eso me le si siñor, delante	This is [given?] to me, master, in face
de la honras de	**of the honors**
Anagoras, siñor, y díceme siñora	And, now, master, lady Clavela tells
Clavela: callan,	me silence [?],
fija Guiomám aprender ben a colar	daughter Guiomán has learned well how to [?]
la flores. . .	**the flowers**. . .

The noun phrases in example (9) show a pattern of marking plurality not found in Afro-Portuguese texts, in which the plural /s/ is not found at the end of nouns. On the other hand, the patterns here are similar to the ones written by Sánchez de Badajoz. That is, the first element of the noun phrase lacks plural /s/ but at the end of the noun /s/ is consistently found. Other examples by the same author are reproduced in (10) and (11).

(10) *Comedia de Tymbria*, (Lipski 2005b: 39; Rueda 1908)

¡Jesú, Jesú, tal decir a **una dueñas tan**	Jesus, Jesus, such words told to **such**
honradas yo la	**honored ladies** [. . .]
Aunque Dios la quiera hacer	Although God would want to confer
merced **a la personas**, no podemo	mercy to **the persons,** we cannot [?]
contigos. . .	
Yo me la sanare a la lumbre de	I will [cure it?] with the light of
mi caras y de **mi ojos**.	**my face** [plural] and **my eyes**

Similar to example (9), plural marking is not found at the end of the determiner, but is present at the end of nouns. It is noteworthy that the noun phrase *mi caras* 'my face' in example (10) has an /s/ at the end of the noun, although the word has obviously a singular referent.

(11) *Comedia de Eufemia*, (Lipski 2005b: 40; Rueda 1908)

¡Jesú! ofréscome la Dios turo poderoso	Jesus! I offer myself to God all powerful
criador **na cielos** é **na tierras**. . .	creator **of the skies** and **the Earth** [plural]
¿Qué te dirá **cuantas siñoras** tengo yo	What will tell you [?] **how many ladies** I have
Guíate **la Celetinas** que guiaba	**the Celestina** guides you, she who guided
la toro enamorados.	**all those in love**

The words in bold presented in example (11) are consistent with the pattern of plural marking found in Rueda's previously mentioned literary texts. An exception is found in the noun phrase *cuantas siñoras* 'how many ladies' where the plural /s/ is found on both elements of the noun phrase. Indeed, given the possible hypercorrect /s/ and distinct

Table 2. Summary of plural marking in Afro-Portuguese and Afro-Hispanic texts (15th-16th centuries)

Patterns of Plural Marking	Afro-Portuguese Texts	Afro-Hispanic Texts
s__s	suas dedos (few similar examples)	los nabos (many similar examples)
s__ø	–	–
#__s	–	dos veces
#__ø	treze pombio_ (many similar examples)	dus ano_ quirados (few similar examples)
ø__s	la_ gentes	la_ flores
ø__ø	muyto negro	–

The # symbol is used here to represent numeral adjective.

plural marking patterns, there is evidence of increased morphological variation of the time. It is important to note that Lope de Rueda was attempting to represent Black speech, and his texts should not be regarded as a representation of his own speech, but rather of the society in which he lived. Regarding this issue Lipski (2005a: 78) states that "Rueda's texts do not give evidence of widespread loss of syllable-final /s/ in sixteenth-century Andalusia, even among Africans, so the hypercorrect [s] which appears in the *habla de negro* is more likely due to Africans' unfamiliarity with Spanish (suffix-based) inflectional systems, and the consequent haphazard use of final /s/, felt to be authentically 'Spanish' without regard for its grammatical function". This statement is supported by the inclusion of the name *la Celetinas* (cf. the correct word is *la Celestina*, the Golden Age figure of a woman who interceded on behalf of lovers). In the text, the word-internal /s/ is deleted but an extra final-word /s/ is added.

What can be learned from the patterns of plural marking present in the aforementioned Afro-Portuguese and Afro-Hispanic texts? Table 2 is a summary of the types of plural marking present in the texts.

Table 2 reveals some overlap between Afro-Portuguese and Afro-Hispanic texts regarding pluralization strategies. Most patterns are shared among dialects. Both Africanized varieties of these Romance languages share the structure of marking the plural on two elements of the noun phrase ('s_s' pattern). However, it should be noted that this pattern is more consistently present in Afro-Hispanic texts.

In Afro-Portuguese texts, only a few such examples were found, and these texts also showed lack of gender concord. As far as the pattern of marking plurality on only the *second* element of the noun phrase ('ø__s' pattern), both dialects displayed that pattern very often. Marking plurality with a numeral followed by a bare noun was very common in Afro-Portuguese texts. In Afro-Hispanic texts, there was only one such example, and it included an added adjective with a plural /s/ at the end of the phrase (cf. 'dos ano *quirados*'). It is interesting to note that only Afro-Hispanic texts showed a pattern of plurality where the /s/ was present in the noun even if it was preceded by a numeral. Regarding the pattern of marking plurality only in the first element of the noun phrase (s_ø), it was initially expected that this would be the primary type of

Table 3. 17th and 18th century texts included in the present study

Afro-Portuguese	Afro-Hispanic
Anonymous Song, ca. 1647 (Lipski 2005b)	– Lope de Vega (1562–1635) *El Mayor Rey de los Reyes* (Vega Carpio 1930) *Vitoria de la Honra* (Vega Carpio 1930) *Madre de la Mejor* (Vega Carpio 1893) *El Negro del Mejor Amo* (Vega Carpio 1930)
Anonymous Poems Poems # 1,2,3 (Hatherly 1990)	– Pedro Calderón de la Barca (1600–1681) *La Sibila de Oriente* (Calderón de la Barca 1682) *La Rabia* (Lobato 1989) *Entremés Séptimo: de los negros de Santo Tomé* (Cotarelo y Mori 1911)

strategy used in Afro-Portuguese texts, considering the parallel pattern with numeral plus bare noun. But in fact, neither variety revealed this pattern of plural marking. Finally, only Afro-Portuguese texts revealed a pattern where plurality could be expressed when no surface /s/ morphemes are present. It should be noted that the glosses are open to interpretation, as far as referentiality and definiteness.

Regardless of the initially perceived similarities, the information presented in Table 2 points to substantial differences between the literary representations of *habla de negro* in these two languages, which may or may not reflect distinctions in the speech of African born speakers of Portuguese and Spanish.

5. The second act: Early 17th and 18th centuries

In the next two sections I will refer to a collection of texts from the 17th and 18th centuries. First, I will present examples of Afro-Portuguese texts and later those from the Afro-Hispanic literary tradition. Table 3 lists the texts that will be analyzed.

5.1 Afro-Portuguese texts (17th to early 18th centuries)

The authors of the Afro-Portuguese literary examples from the 17th and early 18th centuries (cf. Table 3) are not known. Nevertheless, these anonymous poems are worth mentioning for the patterns of plural marking present in them. The poems were probably sung or performed on stage (Lipski 2005a: 59) and were part of a larger corpus of such texts used in musical presentations of the time.

(12) Excerpt of *Anonymous song*, ca. 1647 (Lipski 2005b: 14, University of Coimbra Archives, manuscript #50, folios 18v-23v)

Sã aqui turo zente pleta	There/We are here all Black people
turo zente de Guiné	all people from Guinea
tambor flauta y cassaeta	drum, flute and [?]
y carvave na sua pé	and [?] on your foot (fem.)
[...]	[...]
Nacemo de **hums may donzera**	We were born of (some) **maiden mothers**
huns rey, que minha deuza he	**some kings**, that my God has/got
que ha de forra zente pleta	that has freed/ should free Black people

The stanzas reproduced in example (12) show lack of gender agreement (*sua pé* 'your foot') as well as number agreement in the noun phrases (*huns rey* 'some kings'). The plural noun phrases in bold have plural /s/ at the end of the determiner and lack the plural marker at the end of the other elements of the noun phrase (*hums may_ donzera_*). In spite of this non-standard use of plurality, Lipski (2005a) states that these poems reveal the greater fluency in Portuguese of Africans living in Lisbon.

Hatherly (1990) presented a selection of anonymous poems probably written to the leaders of *cofradías* or religious brotherhoods established in Lisbon. Samples of these texts are given in (13), (14) and (15).

(13) Excerpt of Poem # 1 (Lipski 2005b: 15; Hatherly 1990)

a muyer, e **os fia** sempre está a fazero	the woman and **the sons** are always making
cruzo **nos boca**, senao está sorando	crosses on **our/the mouths**, or are crying

Example (13) shows lack of number as well as gender agreement (*os fia* 'the sons' or 'the children' showing the pattern of only marking the plural on the first element of the noun phrase. It also reveals a pattern of reducing the inflected possessive pronoun *nosso/a(s)* to one invariant form *nos* which is found presently in Cape Verdean Creole and Papiamentu (Baptista 2002; Gilbert 1987). An alternative interpretation is that *nos* simply reveals a lack of gender agreement where the masculine is used instead of the feminine. (cf. standard Portuguese *nas bocas* 'in the mouths').

(14) Excerpt of Poem # 2 (Lipski 2005b: 15; Hatherly 1990)

Zangana a mi; Deus to livra!	[Mad?] with me, God save you!
Vozo **cus coza** di Apolo?	You [have] **the things** [belonging] to Apolo
Mais vare hun dor **dus barriga**	It's better to have a pain **of the bellies** [in the belly]
erso quin qui tem juizo	[and?] who has brain/ is not crazy

Magi si vozo he temozo	but if you are stubborn
dexa **us negro**, us sege dexa	let **the Blacks**, the [?] let them
vai vozo a pe **cus baetinha**	you go by foot with the [?]
porque quin bebe **us cristalo**	because those who drink from **the crystals**
dus fonte do Vabalina	from **the sources** of the Vabalina
ja si sabe, que **us mizélia**	it is already known that **the miseries**
garra dielle como tinha	[it is attached to him?] as a plague

(15) Excerpt of Poem # 3 (Lipski 2005b: 17; Hatherly 1990)

Turo **esses glande prueza**	All **these great feats**
qui vozo tem nomealo	that you have named
nao sá quem qui faz **us honra**	is not known who does **the honors**
ao Sioro Don Brenalo	to master Brenalo

The noun phrases presented in examples (14) and (15) are strikingly similar to those in (13) as far as plural marking is concerned. There is overall consistency in marking the plural only at the determiner, leaving the noun bare of plural markers. In these poems, other phonetic features typical of *habla de negro* of the time can also be found, such as *r/l* interchange (*mizélia* < *misérias* 'miseries' and *vare* < *vale* 'worth') as well as paragogic vowels (*cristalo*< *cristal* 'crystal').

5.2 Afro-Hispanic texts (17th and 18th centuries)

There are many examples of Africanized Spanish texts in the 17th and 18th century, but questions of authenticity and accuracy are usually raised for this period, since the *habla de negro* becomes, by this time, stereotypical or stylistic in nature. In fact, at the turn of the 17th century, Simón de Aguado wrote *Entremés de los negros* (ca. 1602) which showed some of the same features reported in Lope de Rueda's texts, such as paragogic vowels, the copula *sa*(*r*) and *samo*, loss of final /s/ in verbal suffix *-mos*, among others (Lipski 2005a: 82). In addition, the poetry of Góngora (a prominent Golden Age writer) and the plays by Aguado are not discussed here since there were no instances of plural noun phrases found in the available texts. In fact, with the exception of the loss of the final /s/ in the verbal suffix *-mos* > *mo*, there are no cases of syllable-final /s/ deletion in the texts of these authors.

I will refer, instead, to the most prolific writer of *habla de negro*, Lope de Vega and his contemporary, Calderón de la Barca (see Table 3), whose texts include several instances of plural noun phrases. As with the other 17th century authors mentioned before, Lope de Vega's depiction of *habla de negro* is sometimes considered lacking in accuracy. Vega lived for a period in Seville, home to a large number of Blacks, so it is probable that he had first-hand knowledge of Africanized varieties of Spanish (Lipski 2005a). Supposedly, Vega had come into contact with marginalized speech communities, but his cultural and linguistic representations of Blacks reduced them to *buffoons* and were essentially stereotypical. Also according to Lipski (2005a: 85) this character-

istic of Lope de Vega's writing "increased the likelihood that [they] were padded with established linguistic stereotypes". By the same token, Calderón de la Barca, who did not give the same importance to stage dialects as did Lope de Vega, is also considered to be less accurate in his depiction of Africanized Spanish (Lipski 2005a: 87).

Examples (16) through (19) are reproductions of texts by Lope de Vega.

(16) *El mayor rey de los Reyes* (Lipski 2005b: 51; Vega Carpio 1930)

Otro Rey sa de Arabia,	the other king of Arabia
que llamamo Melchioro	whom we call Melchioro
anque negro caballera	although a Black 'gentleman'
samos toros esotros [...]	we are all these others [...]
Esos tres reyes siguiendo	**these three kings** following
una estrella luminossa	a brilliant star
andamo **seiscientas leguas**	we walked **six hundred leagues**
y le parecemo poco...	and it seemed little...

(17) *Vitoria de la honra* (Lipski 2005b: 54; Vega Carpio 1930)

...¡Voto Andioso verrarero	[?] truthful [?]
que sa Sinvilla la reina	that in Seville the queen
de **cuantas civilidades**	of **how many civilities**
turo lo mundo redoa!	All the world [?]
¡Mal **años** para Madrillos	bad **years** for the [?]
[...]	[...]
para Curdoba e Tuledas	to Cordoba and Toledo
Valadulid en Castillas	Valladolid and Castile [hypercorrect /s/]
y en Capalonas, Valencias...	and in Capalona, Valencia [hypercorrect /s/]

Example (16) shows that number agreement is present in noun phrases, even in the case of a numeral adjective followed by a noun (*seiscentas leguas* 'six hundred leagues'). There is also a possible case of hypercorrect /s/ in the case of geographical names, in the case of *Cordoba* and *Toledo*, this last one represented as *Tuledas*. Example (17) presents the same pattern of plural marking, along with cases of final /s/ at the end of proper nouns. Again, it is very probable that the character was referring to topographical names such as Castilla < *Castillas* or Valencia < *Valencias*. If so, these can be construed as examples of hypercorrect /s/.

(18) *Madre de la mejor* (Lipski 2005b: 56; Vega Carpio 1893)

hacia la banda del Congo	towards the margin of the Congo
y sale el siñolón Febo	already goes great master [?] Febo
yena de cabeyo rojo	full of red hair
al mismo punto vimo	from the same point we see
por **montes, vayes** y **sotos**	over **hills, valleys, thickets**

floriro **turo la planta**	flowering **all the plants**
canela clavo oloroso,	cinnamon, aromatic clove
jenjibre, nuece moscada,	ginger, nutmeg,
pimientas y sinamomos	**peppers** and [?]
luego esmaltará **los prados**	soon **the plains** [will be decorated?]
de tanta liria vistoso.	of/with so much lily [?]

A common feature of Africanized speech was the substitution of intervocalic *r* for *d*, as in *tudo* > *turo* 'everything', examples of which can be found in example (18). One could assume that the phrase *turo la planta* refers to a plural referent 'plants' (cf. modern Spanish *todas las plantas* 'all the plants'), since it follows a description of nature with pluralized words. Indeed, most of the examples are of individual pluralized nouns, such as *montes* 'hills' and *pimientas* 'peppers'. In addition, both the determiner and the noun in *los prados* 'the plains' are marked for plural. The noun phrases in this text normally show the standard number agreement. Although not all the noun phrases in the text are reproduced here, they categorically reveal standard noun phrase number agreement.

(19) *El negro del mejor amo*, (Lipski 2005b: 58; Vega Carpio 1930)

Samo jente dilostre, que no ay cosa	We are people [?], for whom there is nothing
que tenga mase lustre que **los neglos**	that is brighter than **the Blacks**
¡Cayan boso, y dejamo noranmalas [?]	You fall, and let us [?]
que alabe Sofonisba **lo moros**,	which [?] Sofonisba **the Moors**
que vivan Dioso que es como un soldano,	who live in God, that is as a soldier [?]
y que no pode ser mase valente	and who cannot be more brave
Alejando **lo mangos** ni Cipoños!...	Going far from **the mangoes** nor [?]

Example (19) shows variable plural agreement at the determiner (*los neglos* but *lo_moros*). In general, however, Spanish standard morphological rules are maintained in these writings. On the other hand, the monomorphemic /s/ changes to onset position with the introduction of a paragogic vowel (note *Dioso* < *Dios*) thus making phonotactic adjustments. There are several ways of adapting phonotactics in a situation of second language acquisition. In the case of more present-day speakers of African languages in contact with Portuguese, for example, coda consonants are a problem.[7] One way of resolving this is adding a paragogic vowel, creating a new syllable with an onset consonant. Another way is deleting the problematic consonant altogether. It seems that in the case of the speech represented here, the former process was chosen, but only selectively, in monomorphemic words. It is also possible that the process of /s/ reduc-

7. See Prata (1983), among others, for a discussion of phonotactic adjustments made by present-day Africans.

tion took place first in high frequency words like *Jesú* < *Jesus*, perhaps from a reduction of *Jesucristo* (Lipski 1995: 141) and later with the plural /s/ of inflected words.

In the second half of the 17th century (in skits written between 1650 and 1670), Pedro Calderón de la Barca also ventured into representations of *bozal* Spanish. Lipski (2005a: 87) has argued that Calderón did not write the most accurate of literary attestations of Africanized Spanish, being restricted to some phonetic modifications and occasionally producing the copular *za* to substitute *ser* and *estar*. Examples (20) and (21) come from Calderón's plays entitled *La Sibila de Oriente* and *la Rabia*.

(20) *La Sibila de Oriente* (Lipski 2005b: 70; Calderón de la Barca 1682)

Az alli rerirara...	There will [?]
Muy pleguntonsica za	very curious/asking she is
No za pozible, que la muzica ze va	it is not possible, that the music goes
y **turos mis gurgunillos**	and **all my groans/sounds**
hasen mucha farta allá...	are very much missed over there

(21) *La Rabia* (Lipski 2005b: 72; Lobato 1989)

Siola, **aquellas seis cajas**	Lady, **those six boxes**
de chocolate me mande	of chocolate send them to me
pagar, pues que las di hasta	pay for them, well I have given already
a **siete reales**, tiniendo	**seven *reales*** [currency], having
tanta parde de Guajaca	so much [?] from Guajaca

Example (20) shows one example of a three-constituent noun phrase *turos mis gurgunillos* < *todos mis gurgunillos* 'all my sounds' which shows categorical plural marking across all elements. Similarly, example (21) presents two noun phrases consisting of a numeral adjective and a noun. In both items, the plural marker is present at the end of the noun. This is the expected configuration if one takes the functional view that all meaningful elements must be preserved. By contrast, it could simply show a tendency to preserve parallel structures in sucessive sentences throughout the discourse (Labov 1994: 550). What can account for the more standardized plural marking in the speech of Black Africans represented here? It is possible that the halting pidgin once spoken by the *bozales* of the Iberian Peninsula no longer existed at the time of these writings. Since literary representations can be unreliable, alternative interpretations are also warranted. This becomes specially true for the last few examples.

Table 4 is a summary of the patterns of plural marking found in 17th and 18th century texts presented in the preceding two sections.

The data summarized in Table 4 is strikingly different from the overall pattern found in previous centuries (refer to Table 2). There is a marked difference between Afro-Portuguese and Afro-Hispanic texts, with pluralization patterns in basic complementary distribution. That is, plural marking in Afro-Portuguese texts follow the pattern of marking the plural solely in the *first* element of the noun phrase, eliminating the /s/ in subsequent elements (*os fia_* ; *huns rey_*). On the other hand, the preferred pattern of plural marking in Afro-Hispanic texts is one of marking in the *second* ele-

Table 4. Summary of plural marking in Afro-Portuguese and Afro-Hispanic texts (17th-18th centuries)

Patterns of Plural Marking	Afro-Portuguese Texts	Afro-Hispanic Texts
s__s	–	cuantas civilidades
s__ø	os fia_	–
#__s	–	esos tres reyes
#__ø	huns rey_	–
ø__s	–	lo_ moros
ø__ø	–	turo la_ planta_

ment, while the first element can be marked or not (*cuantas civilidades, esos reys, lo_ moros*).

In general, the patterns found in Afro-Hispanic texts reveal more standardized speech, at least from the point of view of pluralization patterns. It is unclear if the referent for the noun phrase *turo la planta* is plural, since it could be a modification of the singular nominal phrase *toda la planta*, that is, 'all the plant, the entire plant'. But it is also possible that the zero markers stand for plurals, as in *todas las plantas* 'all the plants'. From the overall patterns seen here, one can surmise that more variation in plural marking strategies is found in Afro-Hispanic texts. It should be noted that it is precisely in these centuries where more variation regarding the representation of Black speech is found. Do these differences prevail in the 19th and 20th centuries as well? The next section is a discussion of attestations of Black speech in Afro-Brazilian texts and Latin American Spanish writings.

6. The third act: Late 17th to early 20th centuries

The objective of this section is to present examples of Afro-Portuguese and Afro-Hispanic texts that were written after the peak of Golden Age literature. The selection is very brief and at the same time includes a large spectrum of time. This is because there are fewer such texts available after the 17th century. In fact, the literary production in the 18th century of such texts is very small. However, there is an increase in the number of texts representing 'africanized' varieties that appear later in the nineteenth century (Lipski 2005a). The examples included here are also from Latin America (represented by the writings of Sor Juana Inés de la Cruz and more recently, Lydia Cabrera), because there was a need to make this transatlantic literary connection in order to allude to the possible African imprint in the Latin American varieties of these languages. Such an imprint can be construed in a less radical perspective, that is, instead of a prior creolization hypothesis that presuposes a full-fledged pan-American creole to have existed in the Caribbean and Brazil, one could contemplate the idea of partial restructuring as explained in Holm (2004). In this view, neither external nor internal conditions exist for full creolization of Portuguese and Spanish, but rather, only

Table 5. Late 18th and 19th century Afro-Portuguese, late 17th century and early 20th century Afro-Hispanic texts

Afro-Portuguese	Afro-Hispanic
– Literatura de Cordel *O preto e o bugio*[...], 1789 (Coelho 1967) *Plonostico culioso*[...], 1819 (Tinhorão 1988)	– Sor Juana Inés de la Cruz (1648–1695) *'Africanized' Villancicos*, 1676 (Cruz 1952) – Lydia Cabrera (1899–1999) *Por qué:... cuentos negros de Cuba* (Cabrera 1972)

a partial restructuring of more vernacular varieties took place. Such restructuring included, in many instances, the plural marking configurations (such as s__ø and #__ø) evidenced in Afro-Portuguese and Afro-Hispanic texts.

Table 5 presents the list of both the Afro-Portuguese and Afro-Hispanic texts that will be analyzed in the next section. These are by no means an exhaustive list of texts, but a selection of texts where examples of plural /s/ variation was found.

6.1 Afro-Portuguese texts in Brazil and Portugal

In a period that ranged from the late 18th century up until the 19th century, there were other attestations of Afro-Portuguese both in Portugal and Brazil. The examples from the late phase of Afro-Portuguese texts presented here are in the form of calendars and pamphlets generally known as *literatura de cordel*. This type of literature published in Portugal in the early 1800's (Tinhorão 1988) refers to short booklets crafted by poets in rural areas which were sold in the street markets. They were displayed by suspending them on a string, thus the name *cordel* (cf. Standard Portuguese *corda* 'string, rope'). Even today, *literatura de cordel* can be found in rural areas of Northeast Brazil, continuing a tradition that started in rural villages of Portugal. Example (23) is one such text: *Plonostico culioso, e lunario pala os anno de 1819* (as cited in Tinhorão 1988: 215). The glosses of the noun phrases in boldface are provided in brackets, and the entire translation appears after the text.

(22) *Plonostico culioso, e lunario pala os anno,* 1819 (Lipski 2005b: 22; Tinhorão 1988)

Aviso ós pubrico: Amado Flegueza, mim vai a continuar com **os Repertoria** [the repertories] dos plesente Anno, e zurgo dever repetir **os Advertencia**, [the admonitions] que **os Repertoria** [the repertories] que tiver **nos Flontespicia** [the frontispiece] ou Subsclita: Porto, na oficina de Viuva Alvarez Ribeiro, e **Filhos ser mia**, [sons are mine; my sons] e outlo quaisquer de Pleto, que appaleça debaixo **dos mia nome** [of+the my names] não sendo ali implesso, ser falso; tomo vozo tento, pala não ser enganaro.

> Notice [?] I make public: Amado Flegueza, I will continue **the repertories** of the present year, and I surmise that I should repeat **the admonitions** that **the repertories** that are in **the frontispiece**, that are underwritten: Porto, in the office of the widow of Alvarez Ribeiro, and the sons, they are **my sons**, and any other Black one, who should appear under **my name(s)**, if they are not written hereunder, are false; I take your [?] so that there is no mistake

The text reproduced in (22) is from an annual collection of pamphlets with predictions of events for the coming year, whose narrators are supposedly Black. The authors used these representations of Black speech as a means of entertainment to white audiences, who recognized such linguistic patterns as belonging to Black speakers. The text in (22) shows some variability regarding plural marking, but the majority of noun phrases carry plural marking in the *first* element (*os Advertencia* 'the admonitions') and no plural /s/ at the end of the noun. There is also considerable lack of gender agreement in the noun phrases, similar to gender marking systems in *Helvécia* (Baxter et al. 1997), an isolated Afro-Brazilian community in Bahia.

In Brazil, literary attestations of *língua de preto* or *língua de guiné* can be found up until the final decades of the 18th century (Lipski 2005a: 60–61). Reproduced here is a fragment of a Brazilian document entitled *O preto, e o bugio ambos no mato discorrendo sobre a arte de ter dinheiro sem ir ao Brasil*, published in 1789 and cited in Coelho (1967: 73–74). Glosses are included in brackets, and the full translation follows the text.

(23) *O preto, e o bugio* [...], 1789 (Lipski 2005b: 24; Coelho 1967)

> Já non pore deixá de incricá **os cabeça** [the heads] e confessá, que vozo doutrina sa **huns doutrina** [some doctrines] tao craro e verdadeiro, que pla mim sá **huns admiraçom** [some admiration(s)] non sé platicada per toro o mundo. O trabaio a que vozo obliga **os pleto,** [the Blacks] e **os blanco** [the Whites] sá **huns trabaio** [some jobs] a que ninguem se pore neá sem melecé **huns cóssa bom** [some good things] porque **os genia** [the demeanors] e **os incrinaçom** [the inclinations] do natureza pore vivé com satisfaçom.
>
> 'Already, I cannot cease to confuse [?] the head(s) and confess that your doctrine is/reflects some doctrines that are so [?] and truthful, that for me, there is some admiration(s) not practiced by all (the world). The work that you force the Blacks to do, and the Whites, are (some) jobs that no one can [deny?] without deserving [?] some good things because the demeanors and the natural inclinations are/exist for one to be able to live with satisfaction'.

In just a few sentences presented in example (23), it is possible to observe many examples of plural noun phrases. In them, the plural is consistently marked in the *first* element of the noun phrase and lacking in subsequent nouns and adjectives (cf. *huns cóssa bom* 'some good things'). Similar to previous Afro-Portuguese texts, the features represented here are lack of gender agreement (*cóssa bom* < *coisas boas* 'good things'), use of *l* in place of *r* (*pleto* < *preto* 'Black'), *r* > *d* change (*toro* < *todo* 'all') and finally, deletion of final *r* of infinitives (*incricá* < *encrencar* 'create confusion'). These are all

features of *habla de negro* consistent with previous literary representations and evidenced in some present-day Portuguese-based creoles, such as Cape Verdean Creole, São Tomense and Papiamentu, specifically lack of gender agreement and deletion of *r* of infinitives (Holm 1988–89; Baxter 1992). In addition, these features are reported for *semi-creoles*, which are contact varieties that underwent only partial restructuring (Holm 2004). These varieties include, according to Holm's definition, Popular Brazilian Portuguese, non-standard dialects of Caribbean Spanish, African American English, Afrikaans, and the Vernacular Lects of Réunionnais French (Holm 2004: xi–xvi). In the case of Popular Brazilian Portuguese, rates of plural /s/ deletion are much higher among less educated speakers (Ferreira 2001), at least in the *nordestino* dialect of Brazil.

6.2 Afro-Hispanic texts in Latin America

Examples of late Afro-Hispanic texts from Latin America come from the work of Sor Juana Inés de la Cruz (1600–1681), who represents a link between the literature of the Iberian Peninsula and Latin America.

From Lipski's collection, I also include here more recent attestations (early 20th century) of Cuban *bozal* speech found in the work of Lydia Cabrera. It should be noted that the work of this renowned Cuban anthropologist is marked by controversy. Cabrera is lauded by some as a popular language pioneer for portraying Cuban *bozal* speech with accuracy. At the same time, she has been criticized by others for negatively exaggerating the speech and culture of the *negros bozales*. Her stories are based on the memories of speech of the Blacks who primarily lived at the end of the 19th to early 20th century. Attestations of Black speech are also found in texts from Peru, Mexico, Colombia and the River Plate region, but due to the more limited scope of this study, these are not included here.

In addition to these caveats, Lipski (2005a: 90–91) states that the Africanized Spanish present in Latin American texts could be based on the speech of Africans living in the New World, but they could also be based on the early literary traditions of *habla de negro* of the Iberian Peninsula. Excerpts of *villancicos* (popular songs performed during Christmas time) written by Sor Juana Inés de la Cruz circa 1676 are presented in example (24).

(24) *'Africanized' Villancicos*, ca. 1676 (Lipski 2005b: 83; Cruz 1952)

Acá tamo tolo	Here we are all
zambio, lela, lela,	[?]
que tambié sabemo	and we also know
cantaye **las Leina**	singing **the queens**
[...]	[...]
Hoy dici que en **las Melcede**	Here it is said that **the Mercedes**
estos Parre Mercenaria	these [for father?]

hace uan fiesta a su Palre,	make a party for her father
[...]	[...]
Camotita linda,	Beautiful *comotita*
fresca requesón,	fresh [?]
qua a **tus manos beya**	that **your beautiful hands**
parece el coló	seems that [?]

Once again, the noun phrases reproduced here show morphosyntactic conditioning of /s/ reduction since plural /s/ is deleted in nouns while the /s/ is maintained at the end of the determiner (*las Leina* < *las reinas* 'the Queens'). Lipski (2005a: 93) states that "this configuration, where /s/ appears only in the first available position of a noun phrase, is typical of vernacular Brazilian Portuguese (also see Guy 1981; Holm 1987), and is found in many basilectal varieties of Latin American Spanish. This is particularly the case with those varieties with a strong African connection, in the Dominican Republic, Colombia, Ecuador and Equatorial Guinea". The second example is from Lydia Cabrera (1899–1999) entitled *Por qué... cuentos negros de Cuba* (Cabrera 1972) and reveals one instance of the pluralization pattern also typically found in the Afro-Portuguese texts.

(25) *Por qué... cuentos negros de Cuba* (Lipski 2005b: 205; Cabrera 1972)

Dende chiquito yo aprende a guerrerar	Since I was a child I learned to fight
Vamo **la siete palma**	let's go to **the seven palms**
lué, ¿quién talla?	It is, who [?]
domiló, domiló	he slept and slept
soto mayimbe nunca duerme.	Only 'mayimbe' never sleeps

In example (25) the numeral adjective *siete* seems to take care of the plural information in the discourse. The common reduction of the first-person plural verb ending *-mos* > *mo*, commonly found in *habla de negro* attestations, in addition to the elimination of prepositions (cf. *vamos* (*a*)... 'we go (to)'), is also evident. No overt plural markers are found at the end of nouns, similar to Afro-Portuguese texts of earlier centuries. There is ambiguity as to the status of the determiner in the phrase *la siete palma*, since the plural could have been present, but the fact that the next consonant in the numeral *siete* 'seven' is also <s> blurs any confident assumptions of categorical /s/ reduction.

What conclusions can be drawn from the later attestations of Afro-Portuguese and Afro-Hispanic texts just presented? First, there are more similarities between Spanish and Portuguese in these texts than in the previous European examples of the two languages. Second, the most frequent pattern of marking plurality is, once again, in the *first* element with deletion in subsequent elements (i.e. *dos mia nome* 'of my names'). Thirdly, there is some variation in the Afro-Hispanic texts, where at least one other element of the noun phrase also is marked for plural (i.e. *tus manos beya* 'your beautiful hands'). Overall, the examples point to a convergence of plural marking in Latin American varieties. One could hypothesize that the mid 19th forced importation of

Blacks to Cuba and Brazil provides the necessary connection between social networks and language use. However, much more is needed to make such categorical claims.

7. Conclusions

The present study uses a compilation of historical Afro-Portuguese and Afro-Hispanic texts gathered by John M. Lipski and traces substantial diachronic evidence of pluralization patterns found in them. The present analyses show that, in general terms, the representations of Black speech in the Afro-Portuguese examples are consistent with a pattern of plural marking independently verified in several comparable systems. More specifically, plural marking patterns point to parallels between well-established Afro-Portuguese creoles (Cape Verdean, São Tomense, Palenquero, and Papiamentu), the Afro-Brazilian dialect of *Helvécia* (Baxter 1992; Baxter & Lucchesi 1999) as well as popular or non-standard varieties of Brazilian Portuguese. Specifically, the absence of plural marking on the head of the noun phrase has often been identified as a vestige of a possible creolizing variety (Castro 1980; Holm 1987) which could have existed in colonial Brazil. The evidence of variation found in earlier Afro-Hispanic texts is less conclusive about the possible contribution of the *bozal* language to modern-day varieties of Caribbean Spanish. As explained in Megenney (2002), all the comparative evidence available today is not enough to state for certain that a pan-Caribbean creole language existed which ultimately influenced Caribbean Spanish. However, the analyses revealed that the texts by Sor Juana Inés de la Cruz as well as the more recent work of Lydia Cabrera show a greater convergence towards the Afro-Portuguese patterns.

Variation in plural marking could also point to second language acquisition patterns, often cited in contrast to evidence of linguistic transfer from African languages (Lipski 1998). It entirely possible that the truth remains somewhat in between. The theoretical construct of partial restructuring put forth by Holm (2004) helps us understand, within a creolist framework, the purpoted African imprint of present-day Latin American varieties of Spanish and Portuguese. The strong preference for the CV syllable and the need for phonotactic adjustments, along with the stark realities of learning a second language in colonial America could have conspired to produce the linguistic characteristics evidenced in these literary texts. Variation in pluralization strategies is only one of several traits which have been cited as part of *língua de preto/habla de negro*; thus, more detailed studies in this area are highly recommended. Aside from more studies that are comparative in nature, it is recommended that parallel quantitative analyses of historical texts be undertaken in order to give a more complete picture of these languages.

References

Azevedo, M. (2005). *Portuguese: A linguistic introduction*. Cambridge: CUP.

Baptista, M. (2002). *The Syntax of Cape Verdean Creole: The Sotavento varieties*. Amsterdam: John Benjamins.

Barrantes, D. V. (Ed.). (1886). *Recopilación en metro del bachiller Diego Sánchez de Badajoz, tomo II*. Madrid: Librería de los Bibliófilos.

Baxter, A. N. (1992). A contribuição das comunidades afro-brasileiras isoladas para o debate sobre a crioulização prévia: um exemplo do estado da Bahia. In E. d'Andrade & A. Kihm (Eds.), *Actas do colóquio sobre crioulos de base lexical portuguesa* (pp. 7–35). Lisbon: Colibri.

Baxter, A. N. & Lucchesi, D. (1999). Un paso más hacia la definición del pasado criollo del dialecto afro-brasileño de Helvécia (Bahia). In K. Zimmerman (Ed.), *Lenguas criollas de base lexical española y portuguesa* (pp. 119–141). Frankfurt: Vervuert.

Baxter, A. N., Lucchesi, D., & Guimarães, M. (1997). Gender agreement as a 'decreolizing' feature of an Afro-Brazilian dialect. *Journal of Pidgin and Creole Languages, 12*, 1–57.

Buendia, F. (1965). *Antología del entremés, desde Lope de Rueda hasta Antonio de Zamora, Siglos XVI y XVII Selección, estudio preliminar y notas*. Madrid: Aguilar.

Cabrera, L. (1972). *Por qué:... cuentos negros de Cuba* (2nd edition). Madrid: Ramos.

Calderón de la Barca, P. (1682). *La sibila de oriente y gran reyna de Saba. Verdadera quinta parte de comedias de Don Pedro Calderón de la Barca*. Madrid: Franciso Sanz. Facsimile ed. by D. W. Cruickshank & J. E. Varey, Vol. XIV of Calderón's Comedias. Westmead: Gregg International Publishers/Tamesis Books, 1973.

Castro, Yeda Pessoa de (1980). Das Línguas Africanas ao Português Brasileiro. *Afro-Ásia, 13*, 81–106.

Chiado, A. R. (1961). *Pratica de oito figuras*. Facsimile ed. by Maria de Lourdes Belchior Pontes. Lisbon: O Mundo do Livro.

Coelho, F. A. (1967). Os dialetos românicos ou neo-latinos na África, Ásia e América. *Boletim da Sociedade de Geografia de Lisboa.* Republished in J. Morais-Barbosa (Ed.), *Estudos lingüísticos crioulos* (pp. 1–233). Lisbon: Academia Internacional de Cultura Portuguesa.

Cossío, J. M. de (Ed.). (1950). *Rodrigo de Reinosa. Santander: Antología de Escritores e Artistas Motanñeses XVI*. Imp. y Enc. de la Librería Moderna.

Cotarelo y Mori, E. (Ed.) (1911). *Colección de entremeses, loas, bailes y mojigangas desde fines del siglo XVI a mediados de XVIII, t.I*. Madrid: Nueva Biblioteca de Autores Españoles.

Cruz, Sor Juana Inés de la (1952). *Obras completas de Sor Juana Inés de la Cruz II: Villancicos y letras sacras.* Mexico: Fondo de Cultura Económica.

Esteves Pereira, João Manuel (1904–1915). *Portugal – Dicionário Histórico, Corográfico, Heráldico, Biográfico, Bibliográfico, Numismático e Artístico.* Volume VI. Lisboa: J. Romano Torres.

Ferreira, F. L P. (2001). Variation in Ibero-Romance: A study of /s/ reduction in Brazilian Portuguese in comparison with Caribbean Spanish. Ph.D dissertation, University of New Mexico.

Garay, R. P. (1988). *Gil Vicente and the Development of the Comedia*. Chapel Hill NC: University of North Carolina Press.

Gilbert, G. (Ed.). (1987). *Pidgin and Creole Languages: Essays in memory of John E. Reinecke*. Honolulu HI: University Press of Hawaii.

Guimarais, A. J. Gonçálvez (Ed.). (1909–1917). *Cancioneiro Geral Garcia de Resende*. Coimbra: Imprensa da Universidade.

Guy, G. (1981). Parallel variability in American dialects of Spanish and Portuguese. In D. Sankoff & H. Cedergren (Eds.), *Variation Omnibus* (pp. 85–96). Edmonton: Linguistic Research.

Hatherly, A. (1990). *A Preciosa de Sóror Maria do Céu: edição actualizada do Códice 3773 da Biblioteca Nacional*. Lisboa: Instituto Nacional de Investigação Científica.

Holm, J. (1987). Creole influence on Popular Brazilian Portuguese. In G. Gilbert (Ed.), *Pidgin and creole languages: Essays in memory of John E. Reinecke* (pp. 406–430). Honolulu HI: University of Hawaii Press.

Holm, J. A. (1988–1989). *Pidgins and Creoles* (2 vols.). Cambridge: CUP.

Holm, J. A. (2004). *Languages in Contact: The partial restructuring of vernaculars*. Cambridge: CUP.

Kiparsky, P. (1982). *Explanation in Phonology*. Dordrecht: Foris.

Lobato, M. L. (Ed.). (1989). *Pedro Calderón de la Barca, teatro cómico breve*. Kassel: Edition Reichenberger.

Labov, W. (1994). *Principles of Linguistic Change: Internal factors*. Oxford: Basil Blackwell.

Lipski, J. M. (1995). Literary 'Africanized' Spanish as a research tool: Dating consonant reduction. *Romance Philology, 49*, 130–167.

Lipski, J. M. (1998). El español 'bozal'. In M. Perl & A. Schwegler (Eds.), *América negra: Panorámica actual de los estudios lingüísticos sobre las variedades criollas y afrohispanas* (pp. 293–327). Frankfurt: Vervuert.

Lipski, J. M. (2005a). *A History of Afro-Hispanic Language: Five centuries/five continents*. Cambridge: CUP.

Lipski, J. M. (2005b). Appendix to *A History of Afro-Hispanic Language: Five centuries/five continents*. <http://www.cambridge.org/0521822653>

Ma, R. & Herasimchuk, R. (1971). The linguistic dimensions of a bilingual neighborhood. In J. A. Fishman, R. L. Cooper, & R. Ma (Eds.), *Bilingualism in the Barrio* (pp. 349–464). Bloomington IN: Indiana University.

Marbán, E. (1971). *El teatro español medieval y del Renacimiento: Una obra para estudiantes de español*. Long Island City N.Y., Las Américas.

Megenney, W. W. (2002). (H)ouve um linguajar crioulo panbrasileiro? *Hispania, 85*, 587–596.

Mello, H. R. de (1996). The Genesis and Development of Brazilian Vernacular Portuguese. Ph.D dissertation, City University of New York.

Pires, S. (1922). Auto da bella menina. In C. Michaëlis de Vasconcellos (Ed.), *Autos portugueses de Gil Vicente y de la escuela vicentina* (pp. 42–46). Madrid: Junta para Ampliación de Estudios e Investigaciones Científicas, Centro de Estudios Históricos.

Pimentel, A. (1889). *Obras do poeta Chiado*. Lisbon: Livraria de Antonio Maria Pereira.

Pimentel, A. (1901). *O poeta Chiado: Novas investigações sobra a sua vida e escriptos*. Lisbon: Empreza da Historia de Portugal.

Prata, A. P. (1983). *A influência da língua portuguesa sobre o suahíli e quatro línguas de Moçambique*. Lisbon: Instituto de Investigação Científica Tropical.

Rueda, L. de (1908). *Obras de Lope de Rueda*. Madrid: Librería de los Sucesores de Hernando.

Seklaoui, D. R. (1986). Parallel Weakening of (s) and Compensatory Change in Italian, French, and Spanish. Ph.D dissertation, University of Michigan. Ann Arbor MI: UMI.

Schwegler, A. (1998). El palenquero. In M. Perl & A. Schwegler (Eds.), *América negra: panorámica actual de los estudios lingüísticos sobre variedades criollas y afrohispanas* (pp. 219–291). Frankfurt: Vervuert.

Tinhorão, J. R. (1988). *Os negros em Portugal: uma presença silenciosa*. Lisbon: Editorial Caminho.

Vega Carpio, L. de (1893). *Obras de Lope de Vega*. Madrid: Real Academia Española.

Vega Carpio, L. de (1930). *Obras completas de Lope de Vega*. Madrid: Real Academia Española.
Vicente, G. (1834). *Obras de Gil Vicente, t. II*. Hamburg: Langhoff.
Vicente, G. (1907). *Obras de Gil Vicente, T.IX*. Coimbra: França Amado.
Vicente, G. (1912). *Obras de Gil Vicente*. Coimbra: França Amado.
Wilkerson, R. (2000). African-American English in film: Be variability. In J. Auger & A. Word-Allbritton (Eds.), *The CVC of Sociolinguistics: Contact, variation, and culture* [Indiana Working Papers in Linguistics 2] (pp. 149–161). Bloomington IN: IULC Publications.

The progressive in the spoken Papiamentu of Aruba[1]

Tara Sanchez
Williams College, USA

This paper investigates the use of the progressive morpheme, *-ndo*, in spoken Aruban Papiamentu using variationist methods. The morpheme is used with durative verbs to mark progressive or repeated action. Achievement verbs with *-ndo* indicate repeated action. *-ndo* use is associated with social prestige, and is used more frequently by each successive generation. These results complement those of Sanchez (2002), which traced the real time development of *-ndo* in its gerundive and progressive uses in texts, and attributed the rise of the progressive function to contact with English. Here we see that Papiamentu *-ndo* behaves more like English *-ing* than Spanish *-ndo*, and that speakers who learned Papiamentu before it came in contact with English do not use this form.

1. Introduction

Papiamentu, an Iberian creole spoken in the Dutch Caribbean islands of Aruba, Bonaire, and Curaçao, productively uses a gerundive/progressive morpheme, *-ndo*, which probably came from Spanish.[2] Since such an inflection-like morpheme is not

1. Portions of this research were supported by a Doctoral Dissertation Research Improvement Grant from the National Science Foundation (#0236758), a Graduate Research Fellowship from the Institute for Research in Cognitive Science at Penn, and a University of Pennsylvania School of Arts and Sciences Dissertation Fellowship. This support is gratefully acknowledged. Thanks to Sarah Bunin Benor, Jeff Conn, Greg Guy, Uri Horesh, Ron Kim, Tony Kroch, Bill Labov, Miriam Meyerhoff, Shana Poplack, Ellen Prince, Gillian Sankoff, Jack Sidnell, John Singler, Kieran Snyder, James Walker, Tonya Wolford, and two anonymous reviewers for helpful comments and suggestions on various stages of this investigation. Any errors are my own.

2. Though Brazilian Portuguese also uses *-ndo*, Portuguese is not thought to exert an influence on its use in Papiamentu since its early removal from the contact situation precludes it from having had a long-term influence like Spanish and English. Since this investigation is based on

commonly a result of creolization (Winford 2003: 322), we must ask whether *-ndo* was productive in the creole from its inception or a later contact-induced development, how it developed in the creole, and how it is used today.

There are no records of Papiamentu at the time of creolization, so we may never know the answer to this first question regarding the status of *-ndo* at this early stage. However, a great deal can be learned about the second, namely the development of this morpheme in the language, by examining early texts and grammars. A detailed analysis of this question based on a diachronic collection of texts is made in Sanchez (2002). I summarize this work below, before returning to a detailed discussion of the third question, i.e. the current use of *-ndo*. It is this last question which constitutes the central topic of the present paper.

Sanchez (2002) illustrates that the morpheme *-ndo* is not found in the earliest texts – a 1775 personal letter and an excerpt from a 1776 court proceeding, both from Curaçao – but it is unclear if this is an accident or an indication that the form was not used in the creole at that time. In these short texts, there are no environments where *-ndo* would be expected. The next available text, an 1803 legal document from Aruba (republished in Maduro 1991: 164), does contain the form (1).[3, 4]

(1) *... y por Ser berdad noos ta firma Ees die noos mismo*
... and be.able to.be true we IMP sign this of our same
mano ***ofres-iendo*** *nos hoeramento delantie die triebunal die nos*
hand offer-GER our oath in.front of tribunal of our
mayor gobierno
greatest governer
'...and for this to be the truth, we sign this with our own hands, offering our oath before the tribunal of our greatest governor'

real time data, and since there are no surviving documents from the Portuguese era, there will be no further consideration of the role of Portuguese in this contact situation. I do this not because Portuguese had no influence on Papiamentu, but because without data from the early period, it is impossible to assess the extent of this influence under the methods used here.

3. The reader will note a great deal of orthographic variation in the Papiamentu examples. Textual examples are given here exactly as they appear in the original. Pre-20th century texts were often written using Dutch orthographic practices. Modern orthography is only partially standardized; Aruba uses an etymological writing system while Curaçao and Bonaire use a phonological one. Spoken examples here are written in the phonological orthography even though all data are from Arubans. This practice is used for the benefit of those unfamiliar with Papiamentu (the sound-symbol correspondence is more regular, making it easier to read) with apologies to Arubans.

4. The following grammatical abbreviations are used: IMP = imperfective, P = past, COP = copula, GER = gerund (refers to translation of *-ndo*), PERF = perfective, COMP = complementizer, 3p = third person singular pronoun, POSS = possessive pronoun, DEF = definite article.

In (1), *-ndo* forms a gerund in an absolutive clause (referred to here as 'gerundive *-ndo*'). A diachronic examination of texts (1803–1916) shows gerundive *-ndo* **only**, and grammars of the time also note only the gerundive use (Evertz 1898; Hoyer 1918; Lenz 1928). During this 100+ year period, the syntax permitted some *-ndo*-marked verbs to surface adjacent to main verb constructions (2) or even the copula (3), (4), but these types were rare.[5] Both (2) and (3) are from translations of the Book of Mark. Spanish and English versions of (3) are also provided to show that the intended interpretation is locative (where *tabata* is a copula[6]), rather than progressive (where *tabata* is the past-imperfective marker (cf. Andersen 1990)). In (4), the copula and gerund surface adjacent to each other, but this is not a periphrastic progressive. Here, too, *tabata* is a copula rather than a TMA marker, and it is part of a relative clause separate from the verb *sclamando* 'screaming'. The intended interpretation is that the husband, not the wife, is screaming. It is not until 1916 that the first periphrastic progressive (referred to here as 'progressive *-ndo*') is attested (5), also from a translation of the Book of Mark.

(2) *…eel a saka soe manoe afoor i a tok-ele* ***bis-ando:*** *mi kieer*
…3p PERF take his hand out and PERF take-it say-GER I want
pa bo bira limpi
COMP you become clean.
'…he extended his hand and took it, saying: I want you to become clean'
Mark 1:41 (1865)

(3) *Juan* ***tabata batiz-ándo*** *den desièrto…*
Juan COP baptize-GER in desert
'John was baptizing in the desert …' Mark 1:4 (1916)
Spanish '*Estaba Juan en el desierto bautizando…*'
English 'John was in the desert baptizing…'

(4) *Ora el a jega ai el a drenta den camber caminda su*
when he PERF arrive there he PERF enter in room where POSS
esposa tabata ***sclamando:*** *Ai mi duci com bo a sufri, bisa*
wife COP exclaim-GER oh my sweetie how you PERF suffer tell
mi kiko a socede?
me what PERF happen
'When he arrived there he entered the room where his wife was, exclaiming: Oh, my sweetie, how you have suffered! Tell me what happened!'
'Carta di Ipi', *La Union* 30 March 1889 (p. 11–12)

5. Though rare, cases of copula (*ta* or *tabata*) adjacent to *-ndo*-marked verbs are significant because on the surface they resemble what became the periphrastic progressive construction: IMP (*ta* or *tabata*) + verb-*ndo*.

6. In Papiamentu, *ta* and *tabata* have dual functions as copula and imperfective markers. *Tabata* is past-marked; *ta* is not marked for time (Andersen 1990).

(5) *...y e Señor* ***tawata trah-ando*** *huntu cu nan...*
...and the man IMP.P work-GER together with them...
'...and Jesus was working together with them...' Mark 16:20 (1916)

Whereas the 1775 and 1776 texts provide no context for gerundive or progressive *-ndo*, the various texts available up to 1916 provide many possible environments for progressive *-ndo,* yet it is not found. I argued that this is because the progressive function was not in use in Papiamentu at that time (Sanchez 2002). The morpheme *-ndo* may have come to Papiamentu at the time of creolization, or later, but either way, only the gerundive function was used at first. After 1930, there was a sharp increase in use of *-ndo* progressives in texts, and grammars from the second half of the 20th century all discuss both gerundive and progressive *-ndo* (Birmingham 1971; Goilo 1953; Howe 1994; Wattman 1953; Maurer 1986; Munteanu 1996). The affix existed in Papiamentu with the gerundive function as early as 1803, and perhaps earlier, but only in the 20th century did the progressive function take root and increase in frequency (Sanchez 2002). This increase coincides with the introduction of English into the contact situation on Aruba. I hypothesized that the Spanish-origin *-ndo* was equated with English progressive *-ing*, leading to the extension of *-ndo* to the progressive function. In other words, gerundive *-ndo* came from Spanish, but the emergence of progressive *-ndo* happened only under the influence of English. And while the various Caribbean Englishes spoken by immigrant oil refinery workers may have had some influence on Papiamentu, the data suggest that the introduction of progressive *-ndo* was a "change from above" (Labov 2001:272–273) and thus more likely influenced by the more or less standard American English of the American refinery managers and engineers, and the variety taught in classes sponsored by the refinery.

This paper investigates how the form is used today, and uses apparent time data to add to the discussion of how it developed in Papiamentu. Specifically, it reports on the linguistic and social constraints on the use of the *-ndo* progressive in the spoken Papiamentu of Aruba, and addresses its status as a contact-induced change from demographic, social, and historical information, and apparent time data. Results support and expand upon the diachronic analysis of texts presented in Sanchez (2002). The data for this paper are taken from a larger set collected during six months of fieldwork in Aruba, Bonaire, and Curaçao in 2003.[7] Here, 43 sociolinguistic interviews (28.9 total hours of speech, averaging 40 minutes each) from a socially stratified sample of Arubans are analyzed using variationist sociolinguistic methods including multivariate analysis.

The paper is organized as follows: First, I describe the language contact situation and provide a detailed sociohistorical background of Aruba in the 20th century. Next, I describe data collection and methods, the variable and linguistic constraints on its use, then discuss its social conditioning. Finally, I place the linguistic and social constraints

7. The sample contains 132 interviews in total: Aruba (n=52), Curaçao (n=53), Bonaire (n=27).

on the variable into sociohistorical context, highlighting how social changes led to this contact-induced change.

2. Language contact in Aruba

Aruba lies just 14 miles from the coast of Venezuela in the Caribbean. At 69.5 square miles, it is the smallest of the three islands where Papiamentu is spoken, and with over 90,000 residents in 2000, it is also the most densely populated. Aruba's population is concentrated in two urban areas – Oranjestad, the capital, in the northwest, and San Nicolas, the refinery town, in the southeast. The areas of Aruba traditionally considered 'rural' lie roughly between the two cities and stretch to the northeast edge of the island. The north coast is uninhabited due to rough terrain, and there is a large, protected, national park area in the southeast, north of San Nicolas and the refinery. Today, the rural areas are turning suburban so that urban sprawl abuts the uninhabitable areas and park land.

Most residents of Aruba use four languages for communicative purposes, albeit with varying degrees of proficiency.[8] Papiamentu is the most common home language, and is widely used in public forums such as religious services, festivals, political speeches, daily newspapers, and local television; Dutch is the official language, and is used as a language of instruction in schools, in government offices, and in official government communication (especially written); Spanish and English are commonly used with tourists and in commerce. They are the languages of popular music, and children learn them in school. Cable and satellite television allow access to American English, Venezuelan Spanish, and European Dutch stations. The four language situation is so normal for Arubans that when asked what language(s) they speak, a typical response is, *Tur kuater* 'all four.'

A growing minority of Aruba residents are not L1 Papiamentu speakers. Many residents of San Nicolas are Caribbean English speakers who immigrated, or whose parents immigrated, to work in the oil refinery. They often speak English at home and learn Papiamentu in the street. Some European Dutch live and work on the island, but because there is not a large community of them, they cannot exist completely in Dutch as they do in Curaçao. They make efforts to communicate in Papiamentu and continue using Dutch at home. Recent Latin American immigrants tend to live in or around Oranjestad. Some are proficient in Papiamentu, others are passively bilingual (they understand the language but respond in Spanish), and still others refuse to learn the creole, focusing on Dutch or English, which they consider more practical in the

8. I do not mean to imply that all Arubans have near-native fluency in all four languages. Most speakers have communicative competency – they can ask for and get what they want, make themselves understood, carry on a conversation, etc. – but some older speakers have limited skills in one or more of the three European languages. I consider both level of fluency and amount of formal education in each of the languages in the quantitative analysis.

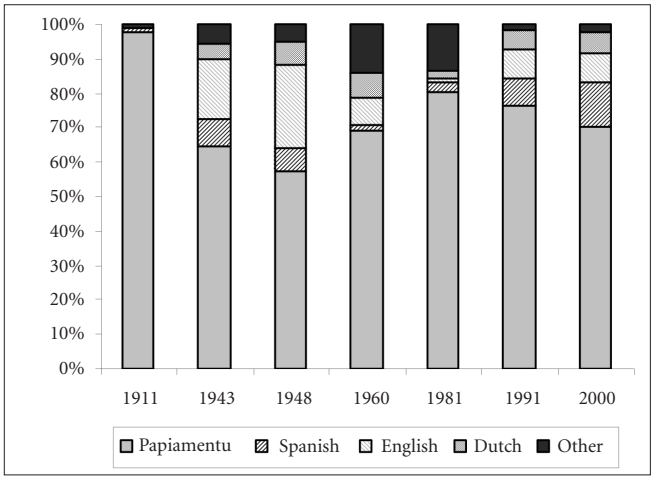


Figure 1. Proportion of Aruban population by 'language used most often at home'[9]

world economy. Though most Arubans can speak Spanish, some refuse to speak it to immigrants on the grounds that those who hope to make their living in Aruba should learn the language of Aruba, which is Papiamentu. This attitude comes after a century of heavy immigration and integration. It is difficult to find Arubans who do not have a parent, grandparent, or other relative from Latin America, other parts of the Caribbean, or Europe, and the very islanders who demand integration through language themselves have immigrants in their families.

Figure 1 shows relative changes in language contact during the 20th century. Table 1 shows figures for total population during the same time period. Together, the figure and table show that despite heavy immigration Papiamentu remains proportionally the language most spoken on Aruba. In absolute numbers, Papiamentu speakers have increased from about 9300 in 1911 to some 63,300 in 2000.

Table 1. Total population of Aruba in the twentieth century

1911	1943	1948	1960	1981	1991	2000
9,616	36,064	47,585	53,199	60,312	66,687	90,505

9. Data for 1911–1960 are estimates of 'language most spoken at home' based on census reports of 'nation of birth'; data for 1981–2000 are based on census reports of 'language most spoken at home' (Alofs & Merkies 2001; Central Bureau of Statistics 2001a; Hawley 1960; Hiss 1943; and Pietersz 1985). 'Other' languages include French, Portuguese, and Chinese, as well as 'language not known or not reported.' Recent censuses report on 'language most spoken at home' (e.g. Central Bureau of Statistics 2001a), which is, in most cases, the L1 of all or most members of the household. Most earlier reports do not discuss language directly, but we can

infer something about language from other data, especially data on immigration and country of origin. To show the comparability of L1 estimates derived from different sources, I present calculations based on 'language most spoken' and 'nation of birth' for Aruba in 2000 in Table A (the only year in which both statistics are available).

Table A. Comparison of 'language most spoken' and 'country of birth' as indicators of L1/dominant language

	Language most spoken	Nation of birth
Papiamentu	69.42%	69.66%
Dutch	6.12%	5.62%
Spanish	13.16%	16.51%
English	8.10%	2.87%
Other language	2.32%	2.93%
Language unknown	0.88%	2.42%
Total Population	90,505 (86,408)*	90,505 (89,990)+

*The total population of Aruba in 2000 was 90,505. The percentages calculated here assume a total population of 86,408 because children under 3 and mentally handicapped persons who cannot speak were excluded from 'language most spoken' data.
+The percentages calculated here are based on 89,990, or the number of residents who indicated a nation of birth.

The second column of the table is an estimate of L1 derived from data concerning the nation of birth. I arrived at these numbers by assuming that the L1 of everyone born on Aruba, Bonaire, and Curaçao is Papiamentu, the L1 of everyone born on Sint Maarten, Sint Eustatius, Saba, Jamaica, the U.S., and England is English, the L1 of everyone born in Suriname or the Netherlands is Dutch, and the L1 of everyone born in Colombia, Venezuela, and Central America is Spanish. Of course, birth in a country does not mean that L1 is the most commonly spoken language in that country. In particular, we know that some people born in the Papiamentu-speaking islands do not speak Papiamentu natively, about 40% of residents of the non-Papiamentu-speaking Antilles (Sint Maarten, Sint Eustatius, and Saba) do not have English as their first language (in 2001, Central Bureau of Statistics 2002a), Suriname is home to Sranan Tongo speakers, and I made the personal acquaintance of Hindi-speaking Surinamese. We can see in Table A, however, that that the two methods make quite similar predictions. Some of the slight differences noted may be due to small numbers of speakers of 'other' languages (e.g. Haitian Creole French, Sranan, etc.), and since a fairly substantial number of speakers are needed for any language to have precipitated contact-induced change, the figures used here should be close enough to reality for the purposes of this paper. The biggest difference in estimates between the two methods is with respect to Spanish and English speakers: the less satisfying 'nation of birth' method clearly **underestimates** the number of English speakers and **overestimates** the number of Spanish speakers. The analysis presented here crucially rests on the strong presence of English speakers in Aruba as indicated by the conservative estimates of the 'nation of birth' data; greater actual numbers of English speakers would lend further support to the argument. When 'nation of birth' is used to estimate L1, it is only because other data is not available.

Another source of error is associated with the specificity of census data – many times, a specific country is not indicated for immigrants, and in these cases, no L1 can be predicted by this

In 1911, an estimated 97% of residents used Papiamentu most often, and at this time there was *not* widespread bilingualism in Dutch and English. Schools operated in Dutch, for the most part, but most students left after only a couple of years because they did not know Dutch before entering school, and could not learn enough Dutch in the early years to make the endeavor worthwhile. Teachers were usually Antilleans who did not have a good grasp of Dutch, but texts, if they existed, were in Dutch and written for native speakers of Dutch. Some religious schools operated in Spanish in the 19th century, and many Arubans spoke this language to some degree at the opening of the 20th century (Green 1974); other religious schools operated in Papiamentu (Catholic) or Dutch (Protestant) (Fouse 2002: 142–143. Hartog 1961: 296). English was practically non-existent on the island at this time.

The Lago Refinery, which operated in English, opened in 1927, providing work for Arubans, other Antilleans, and large numbers of foreign workers, including both Caribbean English-speaking laborers and American English-speaking engineers and managers. Lago sponsored English and other training classes for workers. By 1935, a decree for educational changes was issued because islanders did not possess the minimum skills needed for even the most basic of refinery jobs. Starting in 1906, teachers had been permitted to use Papiamentu when students did not understand Dutch, but the new mandate was that teachers use 'Dutch only.' (Fouse 2002: 144). Though a 'Dutch only' policy was employed in the 19th century to no avail, in the 20th century, teachers were sent from the Netherlands to help improve the educational system, using European teaching materials and standards.[10] Students were now learning Dutch, and math and other subjects through Dutch. In addition, English, Spanish, French, and German classes were introduced.[11] Islanders born in 1930 or later probably encountered the newer educational system.[12] Thus, though the number of Dutch speakers on

method. This usually happens when there are only a handful of immigrants from a particular country or continent, and the census reports say, for example, 'other South American country' or 'African country'. These cases plus cases where immigrants did not indicate a country of birth are the reason percents do not add up to 100. As before, these immigrants make up such a small part of the population that they do not effect the larger argument.

10. Before Lago opened, Royal Dutch Shell opened a refinery on Curaçao in 1915, which operated in Dutch. Many Europeans immigrated to hold engineering and management positions there. They expected their children to obtain an education equivalent to that available in the Netherlands. That, and the fact that the Netherlands agreed to finance primary education in the ABC islands from 1920 (Fouse 2002: 146), drove the educational improvements for Dutch immigrant children as well as Antillean children on Aruba, Bonaire, and Curaçao.

11. In addition to the usual four languages, many speakers between the ages of 40 and 60 also noted that they had studied French and/or German in school, but since they had no opportunity to use these languages in daily life, they had mostly forgotten them. Knowledge of Dutch, English, and Spanish, on the other hand, is reinforced through use in Aruban society and is more easily maintained after schooling is complete.

12. Educational changes were first instituted in Oranjestad, and later spread to rural areas.

Aruba has remained steady and small throughout the 20th century, Arubans' competency in Dutch has increased due to efforts by the government to improve education and the government's insistence on Dutch as a language of instruction.

In 1948, when the refinery peaked in production and number of employees, only about 57% of residents used Papiamentu most often; English speakers comprised about a quarter of the population. Later in the century, as foreign workers were laid off from the refinery and left the island, the proportion of L1 Papiamentu speakers increased again, hovering between about 70 and 80% in the latter half of the century. Interestingly, the children of laid-off foreign workers often chose to remain on the island. They had been born and raised on Aruba, learned Papiamentu from other children, and felt more Aruban than anything else (Green 1974: 107–109).

There were two waves of Spanish immigration in the 20th century – one early on as workers were drawn to the refinery, though their numbers were few compared to English speakers, and another in the latter half of the century to fill construction and cleaning jobs created by the hotel boom. This wave is still in progress, though construction has dropped off sharply, leaving many unable to find work.

Economic changes in the 20th century led to two major social changes: First, there was heavy immigration of non-Papiamentu speakers. Second, a society evolved where four languages are commonly used – Papiamentu, Spanish, English, and Dutch. Immigrant families in Aruba have a tradition of integrating via Papiamentu (Alofs & Merkies 2001; Green 1974). Though there are small enclaves of monolingual immigrants, their children learn the creole quickly, as do some of the immigrants.[13] Multilingualism is an adaptive strategy which allowed Arubans to keep their language. Dutch and English had official and legitimate uses in education and business; Spanish has always been regarded as useful in conducting business with Aruba's South American neighbors. Papiamentu, on the other hand, had no official status for much of the 20th century, was not recognized as a 'real' language, and was completely unknown to the vast majority of immigrants when they arrived on the island. Arubans could have succumbed to pressure to use Papiamentu less in favor of the European languages. Instead, they learned the European languages, but insisted on using Papiamentu at home and with immigrants, thus ensuring the continuity of the language. In the latter part of the 20th century it has gained in status; the language is taught as a subject in secondary schools on the three islands, may be used in the early grades to facilitate the learning of Dutch or in the later grades to facilitate understanding of subject matter, and competence in it may be shown to support citizenship applications on Aruba.

13. The notable exception to the integration norm is the Americans. When working in the refinery, they lived apart from the islanders and only came into contact with them at work. Workers left the island when their contracts were up or upon retirement.

3. Data and methods

To assess the development of progressive *-ndo* with apparent time data and the potential roles of social factors in its development (i.e. in terms of it being a 'change from above'), a socially stratified sample of Arubans was interviewed. Interview subjects were selected opportunistically, either 'cold' or through contacts, to fill cells in a grid stratified according to age, sex, social class, and place of residence. Thirty-nine adult speakers were divided into three age groups: <30, 30–50, 50–70. Four speakers over 70 were also interviewed. Urban or rural residence[14] and social class[15] were defined according to census data and native speaker input. All urban speakers are from Oran-

14. The Noord, Paradera, and Santa Cruz regions are traditionally considered 'rural' by Arubans, but the recent population explosion makes many geozones in these areas almost as populous as Oranjestad. Rather than considering these entire regions rural, I counted only the most sparsely populated geozones as rural. This consisted of five geozones (out of 55 total in Aruba and 48 populated) which were populated but had the lowest population density. These were Alto Vista (Noord); Ayo (Paradera); and Cashero, Urataca, and Balashi/Barcadera (Santa Cruz). Other geozones in Noord, Paradera, and Santa Cruz were considered 'urban' along with Oranjestad East and West (as they are called in the Census). San Nicolas North and South and Savaneta were considered to be part of San Nicolas; no speakers from these areas were included in the sample analyzed here.

15. To assign Arubans to social class groupings, I started with a consideration of the modern appliances, facilities, and electronics a person owned because recent census data (2001a) shows stratification along these lines. If a person did not have electricity or water in their home, they were assigned to the working/lower class (WL) group. A person who answered yes to two points in the WL group in Table B was classified with the WL speakers. A person who answered yes to two points in the upper/ middle class (UM) group was grouped with the UM speakers. A person who does not have two points on either scale (say, has a phone at home, has a TV, has a car, but does not have cable, a cell phone, or a computer) fell in the middle. Additional factors were considered in assigning social class – size of home, income, education, occupation. Decisions were made on a case by case basis.

Table B. Criteria for social class in Aruba

Upper/ Middle Class (UM)	Working/ Lower Class (WL)
computer at home	has no phone in home
internet access at home	has no TV
more than one air-conditioner in home	has no car
cable or satellite TV in home	has no electricity or water
phone in home AND has cell phone	

Table 2. Social stratification of Aruban speakers

Residence	Class	Age	Sex	#	Minutes
Urban	Upper/Middle	<30	F	4	115
			M	3	127
		31–50	F	3	78
			M	2	51
		51–70	F	3	158
			M	3	171
		70+	F	2	134
	Working	<30	F	3	104
			M	2	40
		31–50	F	2	42
			M	2	87
		51–70	F	2	68
			M	0	0
Rural	Upper/Middle	<30	F	0	0
			M	2	64
		31–50	F	3	163
			M	0	0
		51–70	F	1	17
			M	1	27
	Working	<30	F	0	0
			M	0	0
		31–50	F	1	28
			M	1	45
		51–70	F	1	26
			M	0	0
		70+	F	1	120
			M	1	68
Total			26 F	43	28.9 hours

jestad; no San Nicolas residents were included in the sample.[16] Twenty-six speakers were women. Table 2 illustrates the stratification.

16. Since I argue that the use of progressive *-ndo* in Papiamentu was influenced by English, the reader may wonder why San Nicolas speakers were not included here. Time and budget constraints on the project made it impossible to obtain complete stratified samples from both Oranjestad and San Nicolas, and since the larger project investigates several variables, the focus was on the core Papiamentu-speaking community of Oranjestad. To supplement this data, I interviewed seven L1 Papiamentu speakers from San Nicolas. They include one working class male and six middle class men. Several of the middle class men worked in management at the refinery previously. Their data was not included in the VARBRUL run since it is potentially from a different linguistic system. It is discussed in detail in my doctoral dissertation (Sanchez 2005). Here I will simply note that the San Nicolas speakers use progressive *-ndo* with a slightly greater frequency than the Oranjestad speakers.

Only one speaker was not an L1 Papiamentu speaker. She was a woman from the Dominican Republic who emigrated as an adult over 10 years before the interview and has obtained Dutch citizenship. She is an urban, working-class female in her 40s; her L1 is Spanish.

In each interview, I gathered personal information, and then used conversational modules (Labov 1984) adapted for use on Aruba. Topics included language(s) spoken, dialect differences in Papiamentu, and aspects of family, work, island life, religion, politics, as well as other topics determined by the interests of interviewees. I conducted all interviews myself (I am a non-native speaker). Interviews were transcribed, and all tokens of *ta/tabata* + verb (with or without *-ndo*) were counted. The 43 speakers interviewed produced about 9800 tokens; results here come from a subset of 2100 tokens coded for various linguistic factors. The first approximately fifty tokens per interview were used for each speaker. Data (tokens) were analyzed using the GOLDVARB statistical program (Robinson, Lawrence, & Tagliamonte 2001), but specific types were also considered and are discussed here as necessary.

4. Definition of the variable

The variable under investigation is the use or non-use of *-ndo* to express progressive or a related aspect.[17] It is found in root, subordinate, and relative clauses. In these syntactic positions, the imperfective markers *ta* and *tabata* are followed by a verb (6a, c), which optionally takes *-ndo* (6b, d). Differences in interpretation of the forms are discussed in 5.1. The data in (6) are constructed for illustrative purposes.

(6) a. Mi **ta** **bai** skol.
I IMP go school
'I go to school' (single event, progressive, or habitual)

b. Mi **ta** **ba-iendo** skol.
I IMP go-GER school.
'I'm going to school' (progressive or habitual).

c. Mi **tabata** **bai** skol.
I IMP.P go school
'I used to go to school' (past progressive, past habitual, past imperfective)

d. Mi **tabata** **ba-iendo** skol.
I IMP.P go-GER school.
'I used to go to school' or 'I was going to school' (past habitual, past progressive)

17. As in Spanish, this morpheme surfaces as *-ando* or *-iendo*, depending upon the final vowel of the verb.

Table 3. Aspectual interpretation of verb forms

Aspect	Description
stative	state of existence
progressive	action is durative and in progress at time of speaking
repeated	habitual, iterative, or occasional action
imperfective	action is durative but not in progress at time of speaking

In Papiamentu as in Spanish, *-ndo* can also occur in gerunds in absolutive phrases, as in (7), without TMA markers (gerundive *-ndo*). Like Spanish *-ndo* and English *-ing*, the gerundive and progressive forms in Papiamentu are most likely related, but textual evidence suggests a separate, late emergence of the latter. This paper focuses on the emergence and development of progressive *-ndo*, which occurred after the creole stabilized.

(7) *...i* ***kant-ando*** *na bos altu el a bolbe kas.*
...and sing-GER in voice high 3p PERF return home
'...and, singing in a high voice, he returned home'

5. Results

5.1 Linguistic factors

Tokens were coded for aspectual interpretation of the verb and semantics of the predicate. Possible aspectual interpretations are shown in Table 3: stative ((8), (9)), progressive (10), repeated action (11), or other imperfective (12). The verb *ta kanando rond* 'walking around' would not normally be considered stative, but in (8), it is clear from the discourse context that the speaker uses 'walking around' to mean 'existing'. In example (9), the *-ndo* marking indicates a new state. The speaker, a woman in her 50s, was asked if she lives with her mother. She replied that she was, because her mother had had a stroke several years ago, and has been feeling lonely recently. The new state is her mother's loneliness. Conventional stative verbs without a special interpretation were not found marked with *-ndo.* Example (10) was uttered in response to the question, 'How many years of school have you finished?' The speaker wants to know if he should include this year. Example (11) is a response to, 'Are you employed?' In (12), the speaker is discussing how police handle themselves in general. The *-ndo*-marked verb can be considered 'in progress' at the time indicated by the adverbial phrase *den sierto momento* 'at certain times'.

(8) *Bo no kier pa tin un persona* ***ta kan-ando rond*** *ku*
you no want for have a person IMP walk-GER around with
adishon di droga...
addiction of drugs...
'You don't want to have a person walking around with drug addiction...'

(9) *Awor si, pasobra e* ***ta ked-ando*** *muchu su so.*
now yes because 3p IMP keep-GER a.lot POSS alone
'Now, yes, because she's staying by herself a lot'

(10) *E ana aki mi* ***ta kab-ando?***
DEF year here I IMP finish-ger
'This year that I'm finishing (right now)?'

(11) *Mi* ***ta ba-iendo*** *skol.*
I IMP go-GER school
'I'm going to school (as opposed to having a job)'

(12) *Mi ta hana polis den sierto momento* ***ta has-iendo*** *nan best, pero…*
I IMP find police in certain moment IMP do-GER their best but
'I find that the police at times are doing their best, but…'

The semantics of the predicate is analyzed using Vendler's (1967) classification with respect to the features punctual, telic, and dynamic (Table 4). The logical possible combinations of these result in four predicate categories: state, activity (13), accomplishment (14), and achievement (15). A state is mere existence without a goal or change; an activity is some action with duration but no goal; an accomplishment is an action that lasts for some time, proceeding toward and ending at some goal; and an achievement is an action that starts and ends momentarily. Under this system, the entire predicate must be taken into consideration, not just the verb.[18] No true statives were marked with *-ndo* in this corpus. In (13), *papiando* 'speaking' is an example of an activity. In (14), the speaker's husband was very ill for a time, so that *muriendo* 'dying' is meant to indicate a slow process, ending in his eventual death. In (15), the speaker indicates that Arubans repeatedly insert English words into their Papiamentu.

Table 4. Vendlerian paradigm of predicate semantics

	punctual	**telic**	**dynamic**
state	–	–	–
activity	–	–	+
accomplishment	–	+	+
achievement	+	+	+

18. Consider the difference between 'pushing a cart' and 'pushing a cart across the floor'. The prepositional phrase indicates a goal, thus differentiating between 'activity' and 'accomplishment'.

(13) *Ami mes* ***tabata papi-ando*** *hulandes pero mi mamanan tabata papia*
I self IMP-P speak-GER Dutch but my mama-PL IMP-P speak
papiamento ku mi.
Papiamentu with me
'*I* used to speak Dutch, but my parents used to speak Papiamentu with me.'

(14) *Ora k'e* ***tata mur-iendo,*** *nos a yora . . .*
when COMP.3p IMP-P die-GER we PERF cry
'When he was dying, we cried. . . ' (lit. 'Time that he was dying. . . ')

(15) *Nos* ***ta pon-iendo*** *hopi mas palabra. . . vooral na Aruba nos ta*
we IMP put-GER very more word especially in Aruba we IMP
papia mas ingles.
speak more English
'We're putting a lot of words. . . especially in Aruba we speak more English.'

Overall, 92 verb tokens in root, subordinate, or relative clauses were marked with *-ndo* out of a total of 2101. Multivariate analysis picked out semantics and aspect (but not syntactic position) as significant linguistic factor groups. Significant factor groups are shown in Table 5.[19] Here 'weight' refers to the probability of a verb with the given characteristic (progressive, repeated, etc.) to be marked by *-ndo*. For aspect, progressive actions are favored for *-ndo* marking, and repeated actions are slightly favored ((10), (11), respectively), while for semantics, accomplishment and activity ((14), (13) respectively) favor *-ndo*-marking. Statives were disfavored for *-ndo* marking (0.293) under aspectual classification, but slightly favored under the Vendlerian classification (0.611). All stative verbs were stative under both paradigms except two tokens of '*ta kanando rond*' as in (8), which were coded as 'activity' (see also Table 6). Verbs indicating temporary or new states were coded as stative, and it is these verbs which

Table 5. Varbrul weights of aspect and semantic coding, input= 0.016

Aspect	**weight**
Progressive	0.991
Repeated	0.510
Imperfective	0.324
Stative	0.287
Semantics	**weight**
Accomplishment	0.797
State	0.611
Activity	0.573
Achievement	0.315

19. Significant social factors from the same run are discussed in the next section.

Table 6. Cross-tabulation of aspect and semantics factor groups, % *-ndo*

	Other imperfective	Repeated action	Progressive	State	Total
Achievement	– (n=0)	– (n=216)	50% (n=2)	– (n=0)	218
Accomplishment	6% (n= 33)	11% (n=62)	75% (n=20)	– (n=0)	115
Activity	2% (n=249)	6% (n=590)	87% (n=15)	100% (n=2)	856
State	– (n=0)	– (n=0)	– (n=0)	2% (n=500)	500
Total	282	868	37	502	1188

account for the seemingly high weight of statives in the semantics factor group. With these removed, there would be no statives marked by *-ndo.*

There is interaction between these two factor groups; the cross-tabulation is presented in Table 6. The majority of verb forms which were progressive were marked by *-ndo*, as compared to between 0 and 11% of other types of action. The durative nature of 'accomplishment' and 'activity' make it possible for these types of actions to be expressed as 'in progress' with *-ndo* marking, though marked forms may also be, rarely, 'repeated' or 'other imperfective'. The only case where an achievement received *-ndo* marking was where the action was repeated (15);[20] otherwise, the non-durative nature of achievements prevent these actions from being discussed as 'in progress'.[21] So, while achievements are usually disfavored for *-ndo*, repeated (habitual or iterative) achievements may be marked.

Only three speakers did not mark some progressive verbs with *-ndo*; their ages were 82, 80, and 55. The significance of this will be discussed further in Section 6. Four younger speakers (ages 20, 24, and 41 (n=2)) did not use any *-ndo* markings at all. In these cases, no appropriate contexts arose.

5.2 Social factors

Age is a significant factor group (Table 7). Groups up to age 70 have weights over 0.686; the weight for the >70 group is 0.172. Thus, speakers under 70 are significantly more likely to use *-ndo* than those over 70, and each generation is more likely to use this form than the one before it.

---

20. In context, (15) can be considered both a repeated action (English words are repeatedly used) and an action in progress (the repeated use of English words is ongoing now). For coding purposes, one classification had to be chosen (progressive), and this is why there are no *-ndo* marked verbs in the cell 'Repeated action' – 'Achievement' in Table 6.

21. In Spanish and English, achievements do occasionally receive progressive marking if the goal is to emphasize either the unfolding or development of the activity, or the idea that something else (expressed in a separate clause) happened simultaneously with the achievement (e.g. King & Suñer 1980); English achievements may receive a habitual or iterative interpretation (Brinton 1988).

Table 7. Varbrul weights for age, Input 0.016

18–30	*31–50*	*51–70*	*71+*
0.909	0.840	0.686	0.172

Table 8. Varbrul weights for social prestige, input= 0.016

Prestige Group	**Weight**	**Description**
Prestigious	0.919	upper or upper middle class; 'old family'
Neutral	0.670	all others
Not Prestigious	0.327	holds stigmatized job

In addition, the use of *-ndo* is associated with social prestige (Table 8) but not social class. A speaker was considered to have social prestige if a member of the upper or upper middle class, and a member of one of the 'old' families of the island. Speakers holding stigmatized jobs such as domestic help, hotel housekeeping, construction worker, or lottery ticket seller were coded as 'not prestigious'. All others are neutral. 'Social prestige' is an alternate way of looking at social class, but more accurately reflects social categories of importance to islanders. It moves from a two-level distinction to a three-level one, crucially incorporating the notions of the 'old' families and occupational stigmatization (cf. also Green 1974). My reflections on my ethnographic data after I had left the field prompted the development of this coding system. The results are significant since they support the argument (Sanchez 2002) that progressive *-ndo* is a change from above.

Many other factor groups were tested but not selected by the statistical program as significant: place of residence (urban or rural), social class, sex, formality of context (formal speech vs. narrative style), language context of topic (what language does the speaker usually speak when talking about a topic covered in the interview?), frequency and amount of use of languages (Papiamentu, Dutch, Spanish, English), proficiency in the four languages, formal education in the three European languages, subjective reaction to the four languages, and subjective reaction to speakers of the four languages.

6. Discussion and conclusions

The results of this analysis of spoken Papiamentu provide a linguistic description of how *-ndo* is used, and an indication from apparent time data that the use of *-ndo* has changed during the 20th century, and may even function as a sociolinguistic variable.

First let us consider the linguistic (semantic and aspectual) results. We saw that actions marked by *-ndo* tend to be durative actions (i.e. accomplishment or activity)

which are at the same time either in progress or repeated.[22] Achievements are disfavored for receiving *-ndo* marking, but they can take *-ndo* if they are also repeated (habitual or iterative) actions, as in (15). Statives generally are not marked with *-ndo*, but some exceptional cases are. Recall that the stative example (8) contained a normally active verb used to indicate a state,[23] and in (9) the *-ndo*-marked verb indicated a new state. Examination of all the stative *-ndo* cases reveal similar results: there is one other example like (8) from the same speaker, where a non-stative verb is marked by *-ndo* and used to describe a state, two examples of new states, like (9), and three examples of temporary states (i.e. a person temporarily living with parents or staying in a hotel). The *-ndo*–marked imperfective in (12) is in progress at the time indicated by the adverbial clause, rather than the time of speaking. Based on this analysis, we conclude that this morpheme has a coherent grammar in Papiamentu. It is applied productively to Papiamentu verbs,[24] and the linguistic constraints on its use can be clearly delimited.

Next we will consider the significance of social constraints: age and social prestige were significant. That use of *-ndo* is correlated with social prestige suggests that it is a change from above, however, the sample is not evenly stratified by social prestige since this was a post hoc grouping. More data are needed before it can be said with certainty that this is the case.

Differences in the use of *-ndo* based on age are relevant to the morpheme's status as a contact-induced change. In previous work based on analysis of texts (Sanchez 2002), I argued that gerundive *-ndo* existed in Papiamentu for over 100 years, but progressive *-ndo* did not take root until after English was introduced into the contact picture. The results from interviews presented here exactly confirm this finding. Those who learned Papiamentu after the introduction of English in 1927 (i.e. those in the <70 group in 2003[25]) use the *-ndo*-marked form significantly more than those who learned it before English entered the picture (i.e. those over 70 in 2003). Further, though the

---

22. In texts, all *-ndo* forms were progressive (104 marked forms from 5900 total verbs); there were no statives or repeated actions (Sanchez 2002).

23. 'Stative' aspect but 'activity' semantics. An anonymous reviewer argues that this example may instead be interpreted as having habitual, rather than stative, aspect. Such an interpretation also fits with my analysis.

24. Until recently, the affix was only applied to verbs of Iberian origin. Kouwenberg and Murray (1994) report the affix used with the Dutch verbs *zuai* 'swing' and *stofia* 'dust'; I observed *wak* 'watch' → *wakiendo* 'watching.'

25. The age groups 50–70 and 70+ were initially chosen for convenience, not specifically to test the relevance of the introduction of English. The oldest speaker in the 50–70 group is 68; speakers in the 70+ group are 75 (n=2), 80, and 82. Thus, the youngest of these were born in 1928, simultaneous with the opening of the refinery, and others in the group were born before it opened. The oldest speaker in the 50–70 group was born in 1935, the same year that educational changes were announced. By the time children born in 1935 reached school age, these changes should have been implemented.

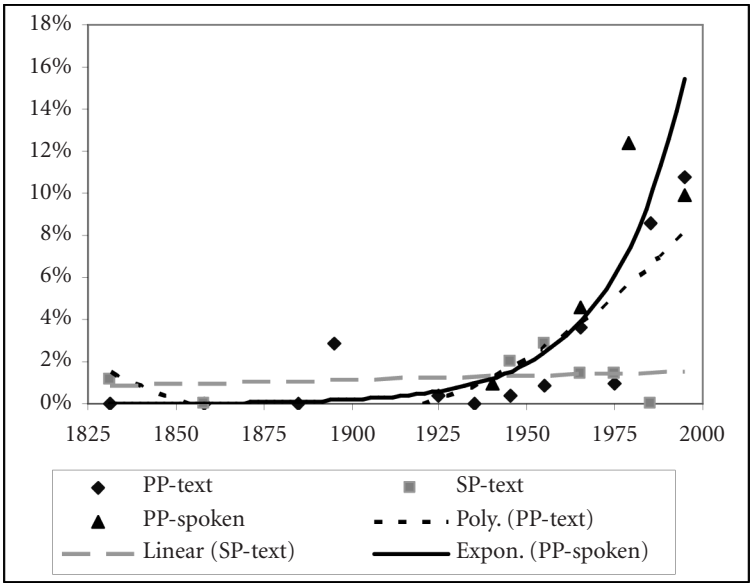


Figure 2. Regression analysis of *-ndo* use in spoken and written Papiamentu and written Venezuelan Spanish[26]

big break is between the <70 and >70 age groups, each generation is more likely to use *-ndo* than the one before it. With respect to linguistic conditioning, the use of *-ndo* with progressive actions is almost categorical; the only speakers (n=3) who failed to mark some progressive actions with *-ndo* were in the two oldest age groups (ages 82, 80, and 55). The latter two findings show that the morpheme is becoming increasingly integrated into the grammar over time, and while it may have been an optional marker in the past, it is an obligatory progressive marker for the younger generations.

Figure 2 compares the increase in *-ndo* in spoken and written Papiamentu to the behavior of the same variable in Venezuelan Spanish texts.[27] While Spanish *-ndo* has not changed much (dashed line), Papiamentu *-ndo* has clearly increased in use in the

26. All interviews were conducted in 2003 and are represented here as apparent time data. Results are reported for each age group at 15 years plus the mid-point of the group. For example, speakers in the 30–50 group were born between 1953 and 1973. The midpoint is 1963, so data is plotted at 15 years after that (1978). In 1978, these speakers would have been between ages 5 and 25. This point reflects an apparent time estimation of how *-ndo* was used by the youngest generation at this time, assuming, as per the apparent time construct, that speakers did not make major changes in their use of the variable later in their lifetimes (Bailey et al. 1991). Papiamentu textual data and Venezuelan Spanish data are from Sanchez (2002).

27. Venezuelan Spanish was chosen because of its geographic proximity to the islands and because there is a long history of contact between the Papiamentu-speaking islands and the mainland which is present-day Venezuela.

20th century in both speech (solid line) and texts (dotted line). Given that Spanish *-ndo* progressives are relatively infrequent, and that Spanish and Papiamentu have been in contact for over 200 years but the change only began about 75 years ago, it seems unlikely that contact with Spanish alone could have caused this increase in the use of progressive *-ndo* in Papiamentu. It is now used at about the same rate as in English (8–10%, Arnaud 1998) and much more frequently than in Venezuelan Spanish (2%, Sanchez 2002).

Previously, I argued that the introduction of English into the contact situation led to these changes (Sanchez 2002). How is this possible given that the form of the progressive, the morpheme *-ndo*, clearly comes from Spanish? English progressives take the form copula + verb + *-ing*, which is analogous to Papiamentu *ta/tabata* (forms of the copula) + verb + *-ndo*. English progressives are frequent (more so than in Spanish), and salient, and thus could cause the frequency of *-ndo* progressives in Papiamentu to increase if speakers relate the two constructions via surface-string matching (Prince 1988; Silva-Corvalán 1993). A similar effect has been reported in cases of Spanish-English contact (e.g. Floyd 1978; Klein 1980).[28]

The linguistic data and additional sociohistorical information presented here show that the opening of the refinery led to the introduction of English, immigration of Caribbean and American English speakers, and educational improvements. The increase in the use of *-ndo* occurs just when we might expect it to if these changes in the linguistic situation precipitated the linguistic change.

Is there any other evidence that English influenced this change, as opposed to Spanish, Dutch, or internal factors? We can easily exclude Dutch from a causal role as it has no analogous progressive morpheme. And while the *-ndo* form came from Spanish, I argue that it is unlikely that the progressive function of Papiamentu *-ndo* came from Spanish since the two languages coexisted (including Spanish's progressive *-ndo*) for a very long time before Papiamentu adopted this function. Additional evidence against the role of Spanish comes from the linguistic conditioning of the variable as determined by quantitative analysis and qualitative reports. The Spanish periphrastic progressive is found with auxiliaries other than *estar* – almost half (47%) of progressives in the UNAM corpus of Mexico City Spanish occurred with auxiliaries such as *ir* 'go', *andar* 'walk, go around', *seguir* 'continue', and *venir* 'come' (Torres Cacoullos 1999). In Papiamentu, *bai* 'go' is used as the progressive auxiliary only 7 times out of 258 *-ndo* progressives from over 9800 verbs in the entire Aruba corpus; none of the other auxiliaries used in Spanish were found in Papiamentu. Further, the four users of this form all have near native proficiency in Spanish and/or close contact with Spanish speakers.[29] Finally, even qualitative accounts of the Spanish progressive present an

28. Spanish-English contact does not always have such an effect, however (cf. Cortés-Torres 2003).

29. One of these speakers received a BA and MA in Colombia. Another was married to a Venezuelan and has a brother who is married to a Venezuelan and lives in Venezuela. The other

analysis which is not consistent with what is found in Papiamentu. For example, King and Suñer (1980: 230) argue that achievement verbs such as *saludar* 'greet' and *dar* 'give' can be marked by *-ndo* to denote developing activity, while in their simple forms they 'imply habitual or future events'. In the Papiamentu corpus, achievement verbs which take *-ndo* can only have a repeated interpretation in context, which is the opposite of what should happen in Spanish. Thus, apart from a few individuals with close contact with Spanish speakers, Papiamentu speakers do not seem to be using *-ndo* as Spanish speakers do.

This leaves the possible role of internal factors to be considered. Progressive aspect developed internally in both Spanish and English, but the process took hundreds of years. In English, the form is attested as early as Tinsdale's 1534 Bible, increasing slowly until the 19th century, when the form quadrupled in frequency (Arnaud 1998). Spanish *estar -ndo* progressives surpassed *ir -ndo* progressives in frequency in 14th century Old Spanish and continue to develop (Torres Cacoullos 1999). While Papiamentu clearly had the linguistic potential for the grammaticalization of a periphrastic progressive with *-ndo*,[30] to say that the development was purely internal would be to suggest that this language developed the form in less than 100 years while others took several centuries. Even if we assume that creole formation is rapid (and recent work indicates that it may not always be (cf. Winford 2003: 306)), the change described here clearly happened after Papiamentu stabilized, and there is no reason to expect purely internally motivated grammaticalization subsequent to creolization (in Papiamentu or any creole) to proceed more rapidly than internally motivated grammaticalization in a non-creole.[31]

The speed with which this development proceeded in Papiamentu suggests that the change was influenced by external factors, specifically language contact. Contact-induced change in stabilized creoles is a fairly well-documented phenomenon. Arends (1989; cited in Winford 2003: 354) describes how Dutch influenced Sranan: Dutch-dominant Sranan speakers import some Dutch features into their Sranan, which are then picked up by younger L1 Sranan (or Sranan-dominant) speakers. In addition, several Caribbean English creoles are undergoing 'decreolization', whereby features of Standard English are incorporated into the creole, moving the creole closer to the stan-

---

two men have a Dominican sister-in-law. Extended families tend to be close in Aruba, so that siblings have frequent social contact with each others' spouses.

30. In other words, Papiamentu permitted the word order copula + verb + *-ndo* with a locative interpretation. The grammaticalization path of this form is reported to be locative → progressive → habitual (Bybee et al. 1994; Torres Cacoullos 1999).

31. An anonymous reviewer argues that grammaticalization in creoles proceeds more rapidly than in non-creole languages since creoles develop in a shorter period than non-creoles. I argue that, once a creole is stabilized, we should expect language change to proceed as rapidly or as slowly as in any other natural language. The change discussed here is one that occurred after Papiamentu stabilized.

dard variety (Winford 2003: 354–355). The only difference between these cases and the Papiamentu case is that Papiamentu is in contact with several languages rather than only one, so we must consider potential influence from each of the languages in the contact situation. In the case of progressive *-ndo*, Spanish contributed the form, but the parameters of its use in Papiamentu do not exactly coincide with those of Spanish.

We are left with (American) English as a potential catalyst for the development of progressive *-ndo*, and there is evidence of this influence in the way the two languages treat achievements, statives, and habituals. Recall that in Papiamentu, the only example of an *-ndo*–marked achievement had a repeated (iterative) interpretation; with respect to English, Brinton (1988: 41) notes: "The effect of the progressive with otherwise punctual verbs is to impart to them an iterative meaning." Brinton's analysis of English *-ing*-marked statives focuses on 'dynamism' (1988: 40), which results in dynamic interpretations of statives (as we saw in (9)), new states (as in (10)), or temporary states (also mentioned above).[32] Finally, Brinton (1988: 40–41) indicates that English *-ing*-marked habituals are interpreted as temporary habits (as in (11)) or hyperbole (as in (15), where the speaker's intention is to show that Arubans use many English words in their Papiamentu).[33]

Aspects of the linguistic conditioning point to English as more of an influence on progressive *-ndo* than Spanish, and coupled with the social conditioning of the variable, the speed of its development, and sociohistorical facts, we conclude that the primary catalyst for the rise of progressive marking in Papiamentu is English.

References

Alofs, L. & Merkies, L. (2001). *Ken ta Arubano? Sociale integratie en natievorming op Aruba, 1924–2001.* Oranjestad, Aruba: VAD/De Wit.

Andersen, R. W. (1990). Papiamentu tense-aspect, with special attention to discourse. In J. Singler (Ed.), *Pidgin and Creole Tense-mood-aspect Systems* [Creole Language Libarary 2] (pp. 59–96). Amsterdam: John Benjamins.

Arnaud, R. (1998). The development of the progressive in 19th century English: A quantitative survey. *Language Variation and Change, 10*, 123–152.

Bailey, G., Wikle T., Tillery, J., & Sand, L. (1991). The apparent time construct. *Language Variation and Change, 3*, 241–264.

Birmingham, J. C. (1971). The Papiamentu Language of Curaçao. Ph.D dissertation, University of Virginia.

32. Statives are disfavored for progressive marking in both English and Spanish (Cortés-Torres 2003; Walker 2001), but see Sanchez (2002) for examples of Papiamentu *-ndo*-marked statives which would be grammatical and felicitous in English but not Spanish.

33. Habituals are disfavored for progressive marking in English and Spanish (Cortés-Torres 2003; Walker 2001), whereas in Papiamentu, repeated actions (habitual and iterative) are favored.

Brinton, L. (1988). *The Development of English Aspectual Systems: Aspectualizers and post-verbal particles.* Cambridge: CUP.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The Evolution of Grammar: Tense, aspect, and modality in the languages of the world.* Chicago IL: University of Chicago Press.

Central Bureau of Statistics (2001a). *Fourth Population and Housing Census, October 14, 2000: Selected tables.* Oranjestad, Aruba: Central Bureau of Statistics.

Central Bureau of Statistics (2001b). *Mapping Census 2000: Social-demographic diversity on Aruba.* Oranjestad, Aruba: Central Bureau of Statistics.

Central Bureau of Statistics (2002a). *Fourth Population and Housing Census: Netherlands Antilles 2001, Volume 1.* Willemstad, Curaçao: Central Bureau of Statistics.

Central Bureau of Statistics (2002b). *Fourth Population and Housing Census: Netherlands Antilles 2001, Volume 2.* Willemstad, Curaçao: Central Bureau of Statistics.

Cortés-Torres, M. (2003). ¿Qué estás haciendo?: La variación de la perífrasis estar + -ndo en el español puertorriqueño. Paper presented at the 7th Hispanic Linguistics Symposium.

Evertsz, N. J. (1898). *Compendio de la Gramatica de Papiamento, ó sea método para aprender á hablarlo y á escribirlo en corto tiempo.* Curazao: Tipografia Bethencourt:

Floyd, M. (1978). Verb Usage in Southwest Spanish: A Review. *The Bilingual Review, 5*, 76–90.

Fouse, G. C. (2002). *The Story of Papiamentu: A study in slavery and language.* Lanham: University Press of America.

Goilo, E. R. (1953). *Gramatica Papiamentu*. Curaçao: Hollandsche Boekhandel.

Green, V. (1974). *Migrants in Aruba: Interethnic migration.* Assen: Van Gorcum.

Hartog, J. (1961). *Aruba: Past and present.* Translated by JA Verleun. Oranjestad, Aruba: D. J. DeWit.

Hawley, A. H. (1960). *The Population of Aruba: A report based on the Census of 1960*. Aruba: s.n.

Hiss, P. H. (1943). *Netherlands America: The Dutch territories in the west.* New York NY: Duell, Sloan and Pearce.

Howe, K. (1994). *Papiamentu Reader.* Kensington MD: Dunwoody Press.

Hoyer, W. M. (1918). *Papiamentoe I su manera di skirbié.* Curaçao: Bethencourt e Hijos.

King, L. D. & Suñer, M. (1980). The meaning of the progressive in Spanish and Portuguese. *Bilingual Review, 7*(3), 222–238.

Klein, F. (1980). A quantitative study of syntactic and pragmatic indications of change in the Spanish of bilinguals in the U.S. In W. Labov (Ed.), *Locating Language in Time and Space* (pp. 69–82). New York NY: Academic Press.

Kouwenberg, S. & Murray, E. (1994). *Papiamentu*. München: Lincom.

Labov, W. (1984). Field methods of the project on linguistic change and variation. In J. Baugh & J. Sherzer (Eds.), *Language in Use: Readings in sociolinguistics* (pp. 28–53). Englewood Cliffs NJ: Prentice Hall.

Labov, W. (2001). *Principles of Linguistic Change: Social factors.* Malden MA: Blackwell.

Lenz, R. (1928). *El Papiamento: Lengua criolla de Curazao.* Santiago, Chile: Balcells and Co.

Maduro, A. J. (1991). *Papiamentu: Indagando i ilustrando.* Curaçao: Drukkerij Scherpenheuvel.

Maurer, P. (1986). *Les modifications temporelles et modales du verbe dans le papiamento de Curaçao.* Hamburg: Helmut Buske.

Munteanu, D. (1996). *El Papiamento, Lengua Criolla Hispánica*. Madrid: Editorial Gredos, S.A.

Pietersz, J. A. (1985). *De Arubaanse Arbeidsmigratie 1890–1930: Drie studies over de trek van arbeiders in het caraibische gebied voor de tweede wereldoorlog.* Leiden: Caraïbische Afdeling Koninklijk Instituut voor Taal-, Land- en Volkenkunde.

Prince, E. (1988). On pragmatic change: The borrowing of discourse function. *Journal of Pragmatics, 12*, 505–518.

Robinson, J., Lawrence, H., & Tagliamonte, S. (2001). GoldVarb 2001: A mulitvariate analysis application for Windows. University of York.

Sanchez, T. (2002). The interacting influences of Spanish and English on the creole Papiamentu. *University of Pennsylvania Working Papers in Linguistics, 8*(3), 235–247.

Sanchez, T. (2005). Constraints on Structural Borrowing in a Multilingual Contact Situation. University of Pennsylvania Doctoral Dissertation.

Silva-Corvalán, C. (1993). On the permeability of grammars: Evidence from Spanish and English contact. In W. Ashby et al. (Eds.), *Linguistic Perspectives on Romance Languages* (pp. 19–43). Amsterdam: John Benjamins.

Torres Cacoullos, R. (1999). Variation and grammaticization in progressives: Spanish -NDO constructions. *Studies in Language, 23*, 25–59.

Vendler, Z. (1967). Verbs and times. In Zeno Vendler (Ed.), *Linguistics in Philosophy* (pp. 97–121). Ithaca NY: Cornell University Press.

Walker, J. A. (2001). Using the past to explain the present: Tense and temporal reference in Early African American English. *Language Variation and Change, 13*, 1–36.

Wattman, F. H. (1953). Papiamentu Morphology and Syntax. MA thesis, Cornell University.

Winford, D. (2003). *An Introduction to Contact Linguistics.* Malden MA: Blackwell.

Was Haitian ever more like French?[1]

Mikael Parkvall
University of Stockholm, Sweden

In the debate on whether or not plantation creoles started out their lives as pidgins, attention has focused on the amount of structure inherited from the lexifier language. Many who argue for a mother-daughter relationship between lexifiers and creoles assume that these similarities derive from the lexifier input in the original contact situation. It has also been suggested that the distance between creoles and their lexifiers grow steadily bigger by time. This paper argues that Haitian Creole, a creole not normally thought of as decreolized, has diachronically moved closer to French. The theoretical implication of this observation is that it allows the possibility of today's Haitian having an ancestor more deviant from French, an ancestor which might have been a pidgin.

1. Introduction

The first creolistics textbook, Hall (1966), had a decisive and longstanding impact on the field. Among many notions which Hall popularised was the concept of decreolization. After pidginization and creolization, Hall suggested that a creole may again become more and more similar to its lexifier, and possibly even be (re)assimilated to it.

In some recent research, it has been suggested not only that creoles fail to develop in the direction of their lexifier, but even that the very opposite is true (e.g., Chaudenson 1979a, 1989, 1992; Mufwene 1992; Alleyne 1996: 175), i.e., that (at least French-lexicon) creoles have diachronically distanced themselves more and more from their lexifier. Also characteristic of this approach is the belief that creoles are not descended from pidgins. Instead, creoles are fully genetic (i.e., perfectly compatible with the family-tree model) descendants of their lexifiers, whose evolution has been accelerated (and otherwise only slightly affected) by processes of second language acquisition.

1. Philip Baker, John McWhorter, Stéphane Goyette, Anthony Grant and Päivi Juvonen provided useful comments on an earlier incarnation of this paper.

A particularly strong stance has been taken in various writings by DeGraff (e.g., 2001a, 2001b, 2001c). One of the main arguments used in favour of his position is that the French-lexicon creole of Haiti – the one creole discussed by DeGraff – has more affixation than one would expect from an erstwhile pidgin. For those who do believe that creoles – including Haitian – are indeed born out of pidgins, it is tempting to assume that some or most of the Haitian affixes are in fact not directly inherited from the lexifier, but have entered the language through later contact with French. The latter possibility is firmly rejected by DeGraff (2001a: 31, 2001c: 22), on grounds that will be discussed later. In other words, like Chaudenson, Mufwene and Alleyne, he believes that Haitian has not drifted in the direction of French since its inception in the mid- to late 17th century.

In this paper, I will show that while definitive proof for either position may be impossible to produce, there are certain data which are very suggestive of French having influenced Haitian after creolization.

2. Geographical distribution of linguistic features within Haiti

One source of such data is Dominique Fattier's (1998) dialect atlas of Haitian. Building on an extensive survey of 20 Haitian localities, it contains more than 2 000 maps. The atlas gives a unique insight into variation within the language, and its quality and value appears uncontroversial in this context since some of those with whom I am taking issue here have repeatedly spoken highly of it (e.g. DeGraff 2002: 347).

Examining the atlas, I recorded about 500 linguistic features (of which 330 were used for the calculations below) for which significant geographical variation was documented. Since Fattier's work is very much a traditional dialect atlas, most of the data are lexical in nature. The lexical items are phonetically transcribed, and thus also allow for phonological conclusions, but morphological and syntactical traits are less easily extracted.

For many of the items considered, there is one more and one less 'French-like' way of expressing things in Haitian. Some localities, for instance, make use of front rounded vowels in etymologically appropriate positions, which is clearly more French-like than not having such vowels. Some varieties also have apparent remnants of French gender agreement, as in *jalou* vs. *jalouz*, both meaning 'jealous', but the first form being used of men, and the other of women (< French *jaloux* ~ *jalouse*). A dialect which has this distinction is closer to French with regard to this particular feature than one which does not.

Using two phonological features typical of French, but not of Haitian (front rounded vowels and post-vocalic /r/ – again in etymologically motivated positions) and 69 individual lexemes of the type just illustrated (i.e., where there is geographical variation between two expressions, one more similar to the lexifier, and another more distant from it), it is possible to develop a relative index of 'Frenchness' for the local-

Table 1. Haitian settlements ranked according to 'Frenchness'. The higher the value, the more 'French-like' the variety

1	Fort-Liberté	18,67	10	Dessalines	9,67
2	Torbeck	18,00	11	Bainet	9,33
3	Saint-Raphaël	16,67	12	Marigot	9,00
4	Limbé	15,67	13	Saint-Louis-du-Nord	8,33
5	Béraud	15,00	14	Aquin	5,33
6	Hinche	12,67	15	Saint-Marc	5,33
7	Beaumont	11,67	16	Dame-Marie	5,00
8	Jean-Rabel	10,67	17	Petit-Goâve	2,67
9	Ganthier	10,33			

ities investigated by Fattier and her assistants.[2] Table 1 below illustrates the ranking. The speech of Fort-Liberté in the north-eastern part of the country is, in this metric, that which is most similar to French, while that of Petit-Goâve is least so.

However, the data presented in Table 1 does not help us to assess whether the dialect of Fort-Liberté shows a high frequency of 'French-like' variants because it has been Gallicised subsequently, or because French never was so thoroughly creolized in the north-east as it was in the centre of the country. To address this issue, we need another index – one of 'archaicness'. A lexifier-like feature which has survived creolization would reflect 17th century (dialectal) French, while we would expect a post-formative influence to more resemble 19th or 20th century (and more standard-like) French.[3]

2. The twenty localities were ranked for each of the three main parameters (front rounded vowels, postvocalic /r/, and the 69 lexemes). For each parameter, the score of a given locality is the number of "French-like" realisations out of the total possible instances, e.g., the number of words with postvocalic /r/ vis-à-vis the number of words (present in the dialect) which have postvocalic /r/ in French. The index figures given in the table here reflect the average position of a location on these three lists. It should be noted that the three parameters correlate with one another, so that using only one of them would have yielded a rather similar result. This strongly suggests that variation on the ±French axis is not just random. Three of the twenty locations featured in Fattier's atlas (Bonga, Cazale and Léogâne) stand out in that information is lacking more often than not, and they are therefore excluded from the analysis.

3. In many cases, of course, there is no difference between 17th and 20th century French. The features classified as "archaic" (i.e., typical of older French) for the purposes of this analysis are those which are not found in modern French. Also, one of the anonymous peer reviewers commented that inherited traits need no longer reflect 17th century French because of internal evolution within Haitian. While this is true, it would be surprising if such a development were consistently in the direction of modern (standard) French – lest, of course, the latter had exerted influence.

Table 2. Haitian settlements ranked according to 'archaicness'. A high value indicates a higher similarity to 17th/18th century French

1	Ganthier	16,33	10	Aquin	8,67
2	Petit-Goâve	16,00	11	Beaumont	8,33
3	Saint-Marc	15,00	12	Marigot	8,00
4	Bainet	10,67	13	Dame-Marie	7,67
5	Jean-Rabel	10,33	14	Béraud	6,00
6	Limbé	10,00	15	Saint-Raphaël	6,00
7	Hinche	9,33	16	Fort-Liberté	5,00
8	Saint-Louis-du-Nord	9,33	17	Torbeck	4,00
9	Dessalines	9,00			

Two well-known phonological changes[4] which took place in French in this period were 1) the loss of /h/, and 2) the development of <oi> into /wa/.[5] For both features, both the older and the more modern version are found in Haitian. Just like in the calculation of the Frenchness index, we may in addition to this use a number of individual words – in this case, I identified 33 suitable items. Examples include *août* 'August' pronounced as /au(t)/, rather than, as in modern French /u(t)/, the retention of final /t/ in words such as *doigt* 'finger' and *nuit* 'night', the use of *estomac* for 'chest', rather than for 'stomach', and archaic dialectalisms such as *icitte* 'here', *chesse* 'dry' and *sectembre* 'September' (modern standard French *ici*, *sèche* and *septembre*). These are not only attested in older and dialectal versions of the lexifier, but were demonstrably taken overseas by 17th century colonists, since they are still found in various French dialects in Acadia, Canada proper, Louisiana, the Caribbean and the Indian Ocean (Highfield 1979; Chaudenson 1979b; Carayol & Chaudenson 1984, 1989; Hollyman 1964; Dorrance 1935; Thogmartin 1979; Griolet 1986; Massignon 1962; Société 1930, among many other sources).[6] As with the Frenchness index, the phonological and lexical evidence correlate with one another, suggesting once again that the variation is systematic. Following the same procedure as in Table 1, the results are as follows:

4. Some details on the relevant changes can be found in any textbook on the history of French, such as Posner (1997: 258–261).

5. The most common earlier pronunciation was /wɛ/, though there was a considerable amount of variation, including the /wa/ which was later generalised. This implies that if an item contains /wa/, there is no guarantee that it is a recent introduction. The opposite generalisation holds, however, so that /wɛ/ does indicate a certain antiquity. It should be borne in mind that the <oi> part of the "archaicness" index is not based on individual occurrences of one or the other, but rather the *proportion* of /wa/ to /wɛ/ (or any other non-/wa/ realisations) in each location.

6. The fact that a given feature was exported from France to various overseas colonies does not of course prove that it was also taken to Haiti. Again, however, it would be surprising if local developments of what was originally standard French input in Haiti happened to yield for case after case forms also found in rural France.

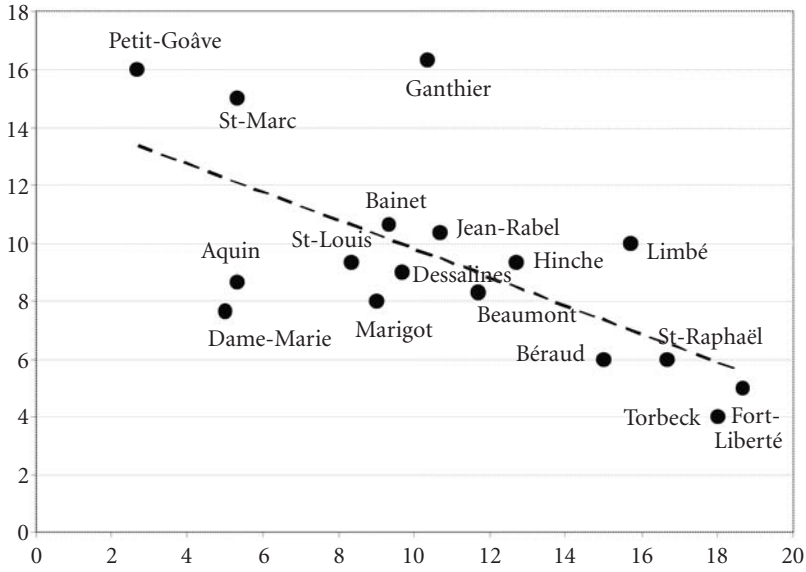


Figure 1. The relationship between 'Frenchness' (X axis) and 'Archaicness' (Y axis) in 17 Haitian dialects

What, then, is the relation between 'Frenchness' and 'Archaicness'? Figure 1 displays the inverse relation that holds between 'Frenchness' and 'Archaicness'. The correlation (–0,65; Pearson correlation) is negative and significant at 0,01 level.

In other words, the locations whose creole is most similar to French are also those where the creole is most similar to *modern* French. The dialects which are in general the least French-like, on the other hand, are those which best preserve 17th century features absent from modern standard French. This suggests that many of the French-like features under consideration in Haitian are in fact not inherited, but acquired at a later date. In the scenario proposed by those with whom I am taking issue here, we would expect the arguably more archaic speech of Petit-Goâve to be more French-like than that of Fort-Liberté.

3. Agglutination in Haitian

The same point can also be demonstrated in another way. Haitian and other French-lexicon creoles are known to display a fossilisation process normally termed agglutination in the creolist literature.[7] This refers to the merger of lexifier determiners (and in some cases also other material) with the noun. Typical examples include *lalin*

7. Please note that the traditional use of this term in creolistics differs from its use in typological, and indeed most linguistic, literature.

'moon' (< French *la lune* 'the moon') and *lanmè* 'sea' (< *la mer* 'the sea').[8] While obviously derived from a article+noun combination, these words are synchronically speaking monomorphemic in Haitian, and may be pluralised or equipped with various determiners without any change to the etymologial article.

Some agglutinated items have incorporated an entire syllable (such as the ones we have just seen), while others have merely received an additional initial consonant. Examples of the latter category are *lò* 'gold' (< *l'or* 'the gold'), *ze* 'egg' (< *des œufs* '(some) eggs' or *les œufs* 'the eggs').

The segmental agglutination type is not unheard of in French itself – more or less rural varieties of Canadian and Louisianan French, for instance, sporadically contain items such as *lécole* 'school' (< *l'école* 'the school') and *nomme* 'man' (< *un homme* 'a man'), and even a few standard French items, such as *lendemain* 'tomorrow', *tante* 'aunt' and *orange* 'orange' display direct or indirect results of such morpheme boundary reinterpretations, as do the oft-cited *newt*, *nickname, apron* and *adder* (< *ewt*, *ekename, napron* and *nadder*) in English (e.g., Anttila 1972: 93–94). Yet, the number of such reinterpretations is extremely limited in either English or French (hence the repetition of the same few items in the literature), while being rather numerous in French-lexicon creoles. What is more, the syllabic agglutination exemplified by *lalin* and *lanmè* above is virtually unheard of in French.[9] Even in a semi-creole such as that of Réunion island in the Indian Ocean, the total number of documented cases is a mere 12 (Grant 1995: 156). On the neighbouring Mauritius and the Seychelles, where dialects of a more thoroughly restructured French-lexicon creole are spoken, Grant (1995: 156) counted well over 600. The same author found 166 in Haitian, to which I could add about 20 more.

The fact that syllabic agglutination is virtually absent from any variety of French underlines the already obvious fact that it results in a rather drastic departure from the basic makeup of this language. If creolization is based on a kind of second language acquisition, which I doubt,[10] but which is the most widely upheld hypothesis, then the learner saying *lalin* for 'moon' has seriously misinterpreted one of the most basic aspects of French grammar.

---

8. About half of these also have non-agglutinated variants. For the present purposes, an item is considered agglutinated if it has such a form, regardless of whether or not a "bare" variant co-exists with it.

9. Canadian French has *mononcle* 'uncle' and *matante* 'aunt' (Stéphane Goyette, p. c.), but then, kinship terms behave differently from other nouns in many languages. In any case, the two are not far from constituting an exhaustive list of such items in Canadian French.

10. This term is inappropriate since "acquisition" implies that people sought to acquire something already existent. In my view, the main interest among those creating a pidgin (which, in turn, I see as a prerequisite for creolization) is mutual understanding. Whether or not this conforms to an already existing language or not would have been of secondary importance (see also Baker 1990 and Parkvall 2001).

Note also that while it is certainly not uncommon for French nouns to be borrowed into other languages in their agglutinated form, agglutination is not particularly characteristic of L2 French (Parkvall 2001). Meanwhile, it is rather common in pidgins lexically based on, e.g., Arabic, Chinook, Fijian, Hindustani, Inuktitut, Swahili and Unami (Goddard 1996: 65; Grant 1996: 235; Kaye & Tosco 1993: 290; Prokosch 1986: 99; Siegel 1987; Thomason 1980: 171; van der Voort 1997: 376–377; Vitale 1980: 51), and is attested also in the few documented French-lexicon pidgins (e.g., Anon 1916).

So what might this have to do with the issue of decreolization (or at least post-formative French influences) in Haitian?

An important observation, I think, is that the agglutinated nouns are not evenly distributed across the language. Recall the two main alternative historical scenarios – on the one hand, there is the suggestion that Haitian has grown out of a pidgin with a limited lexical inventory. A pidgin is an emergency language created among people who for the most part do not speak the lexifier, and who may not even want to do so. The original creators of Haitian, according to this proposal, did not master (and probably did not even seek to master) the French determiner system. In this scenario, then, agglutination would be a feature mainly associated with the pidgin phase. The growth of this hypothesised pidgin into the full-fledged language it is today would have included – among other things – borrowing from the lexifier. Borrowings tend to be channelled into a language through a group of more competent bilinguals. As a result, agglutination in general, and whole-syllable agglutination in particular, would be confined mainly to the basic vocabulary, which is inherited from the pidgin stage. I shall refer to this as the *broken transmission hypothesis*. In the alternative scenario, Haitian is as much a legitimate daughter of French as French itself is of Latin, and at no time in the language's history would there have been a pidgin 'bottleneck' – it has passed from one full-blown state to another. This will be referred to as the *continuity hypothesis*. In this scenario, syllabic agglutination would be unexpected in the first place, but if present (as it obviously is), there would be no reason to expect it to be concentrated in any particular part of the lexicon.

To test the two hypotheses, I devised a Haitian wordlist of 5 400 Haitian words, divided into three categories: *Inner Core, Outer Core* and *Periphery* (see Appendix for definitions). It turns out that agglutination (of both types) affects 10,7% of the Inner Core, but only 4,5% of the Periphery. This is not the whole story, however. Given that nouns constitute the most open class in any language, this category is likely to be better represented the further away from the core that we stray. Indeed, nouns make up 45% of the Inner Core, but 57% of the Periphery. In other words, the items which have had a chance to undergo agglutination, i.e., nouns, are more likely to turn up at the outer fringes of the vocabulary. Therefore, it would make more sense to only count the share of agglutinated items among nouns. This increases the numbers, but on the other hand, I have here ignored proper names such as *Laswis* 'Switzerland' (< *La Suisse*) and *Lazi* 'Asia' (< *l'Asie*). The proportion of agglutination in relation to the total number of nouns is displayed in Table 3 below.

Table 3. Agglutination in different parts of the Haitian lexicon

Agglutination	Inner core	Outer core	Periphery	Total no. of nouns
None	78%	87%	92%	2650
Segmental	11%	6%	4%	150
Syllabic	11%	7%	3%	140

As can be seen, agglutination is considerably more common in the basic lexicon than in the more peripheral parts of the vocabulary. Note also that the Outer Core places itself in between the Inner Core and the Periphery, which is encouraging. If we had only a core and a periphery, a higher percentage in one than in another might be due to mere chance. Now that we take three parts of the lexicon into account, we can see that there is indeed a cline from the most to the least basic items. (Among the at least slightly similar examples found in English, it is striking that the four most oft-cited examples – *adder, newt, nickname* and *apron* hardly form part of what could be called 'core vocabulary'.)

It is well known (e.g., Hock 1991: 385 and Zengel 1962: 138) that the core vocabulary is more resistant to change than the periphery. The Haitian data also provides explicit support of this if we again consider the varying reflexes of French <oi> in Haitian. In the inner core, 69% of the items containing an etymological <oi> are reflected as /wa/, in the outer core this rises to 80%, and in the periphery, this figure is 83%. This indicates, once again, that the core, where the highest agglutination rate is found, is indeed the oldest part of the Haitian lexicon. In other words, the periphery, the part which is most likely to have been affected by outside influence is not only more French-like, but also less archaic. Presumably, no one would doubt that a word for, say, 'fire' would have existed in Haitian before a word for 'telecommunications' did. However, under the hypothesis claiming that the creole has strayed further and further away from the lexifier, there is no reason to expect these to be *dife* (< *du feu*) and *telekominikasyon* (< *télécommunication*) – which represents the actual state of affairs – rather than **fe* and **latelekominikasyon* respectively. The broken transmission hypothesis provides a ready answer: a word for 'fire' was present in what was to become Haitian at a time when the language was still a pidgin, while the word for 'telecommunications' was borrowed from French at a later date. Needless to say, no one would claim that 17th century Haitian had a well developed IT terminology, but the point remains: there is a difference between core and peripheral vocabulary in Haitian which is predictable if we assume pidginization to have played a role in the development of the language, but which is not accounted for by the continuity hypothesis. Note that I am not arguing that agglutination as such is diagnostic of pidginhood – at least sporadically, such reinterpretation may occur in any language. You need not even believe that pidginization is more likely to include morpheme boundary reinterpretations than borrowing is. Instead, the most important observation is this: Had Haitian distanced itself more and more from French, we would expect these changes to first affect the periphery, which tends to be less stable. Therefore, under the scenario I am arguing

against here, we would expect the core to be more like French than the periphery. Yet, upon closer examination, the reverse is true. This does not prove a pidgin past of Haitian, but such a hypothesis accounts for the facts just observed, while the continuity hypothesis simply has no answer.

4. Lexicon vs. structure

So much for the lexicon – the results this far suggest that this part of the language has indeed been subject to French influences after creolization. But what about grammar? After all, the difference between a lexifier and a creole is normally taken to be found mostly in the structure rather than in the lexicon. Morphology is particularly important here, since a drastic loss of inflections constitutes the most salient and conspicuous result of creolization.

DeGraff (2001a: 9–10, 2001b: 7–8) argues for the existence of 17 morphological processes in Haitian which would illustrate that the language is less analytical than than often proposed in the literature, and that not only lexicon but also quite a bit of morphology has been inherited from French. Of these 17, two involve the same morpheme (verbalising *-e*),[11] while another (compounding) is a process so commonly attested in pidgins that it presents no problem for the broken transmission hypothesis. Three other morphological processes – reduplication, and the affixes *ti-* (diminutive) and *-adò* (agentive) may well be present in Haitian, but are not inherited from the lexifier. The agentive is borrowed from Spanish, reduplication is certainly not a morphological property of French,[12] and the diminutive appears to have been grammaticalised within Haitian (from *piti* 'small'). Needless to say, after nativisation, a creole functions just like any other language, and it is not unusual for a language to develop new affixes in 350 years.

That leaves us with 12 Haitian affixes which do have counterparts in French, and which may thus be suspected to have been inherited. This would indeed seem to represent an affix inventory larger than that of a typical pidgin.

---

11. Even if verbalising *-e* and the opposition between so-called long and short forms of the verb are counted separately, that still leaves DeGraff's account with a problem. Apart from the semi-creole of Réunion, this opposition (seemingly modelled on French verb forms such as *mange* ~ *manger*) is attested in three French-lexicon creoles, namely those of Louisiana, Mauritius and Haiti. As it happens, there is abundant evidence from the two former ones that the distinction is a post-crystallisation development (Klingler 2000: 29; Baker & Corne 1982; Corne 1999: 111, 167; Neumann 1985: 197), something which should at least cast some doubt on the assertion that it is inherited in the case of Haiti.

12. In this particular case, I am not convinced that reduplication, as opposed to *iteration* or *repetition,* (for discussion of these concepts with particular reference to contact languages, see my own and other papers in Kouwenberg [ed.] 2003) is a property of Haitian either, but that is of lesser importance here.

However, the most common words with the nominalising *-syon* suffix (< French *-tion*) in my corpus are *alfabetizasyon* 'alphabetisation', *koripsyon* 'corruption' and *konstitisyon* 'constitution'. Regardless of one's views on the role of pidginization in creole formation, it does not seem particularly likely that these were among the first items acquired by 17th century slaves. About six of the remaining affixes discussed by DeGraff are such that they are found in what appear to be local coinages, that is affix-root combinations not attested in French (Valdman 1978: 131–143; Lefebvre 1998: 310–311), something that constitutes irrefutable proof of their productivity in Haitian.[13] This is not the case, so far as I can tell, for the others.

Despite the reservations above, it is clear that Haitian – while clearly a highly analytical language – is not by any means devoid of bound morphemes. It certainly displays more morphology than we would expect from a pidgin. Haitian is of course not a pidgin, but, if it once was, it would require that many of the affixes post-date the formation of the language.

Because of the nature of the data available to me, I have hitherto mostly discussed lexicon and phonology. Nevertheless, lexical loans are not irrelevant to a discussion on structural influence. It is usually believed that when a bound morpheme is borrowed into a language, this is done by the way of lexical borrowing. When a sufficient number of morphologically complex (from the donor language's perspective) words have entered a language, its speakers have acquired a pattern which they can generalise and (in some cases) apply to indigenous items (e.g., Hock 1991: 382). Tagalog, for instance, which used to be a language with neither biological nor grammatical gender, has borrowed Spanish adjectives along with their gender agreement. Indicating the gender of animates has thereby become an obligatory part of the system of Tagalog, and this, in turn, can be argued to have created a hitherto non-existent category of adjectives in the language (Stolz 2002: 146–149). Developments in Ilocano are almost identical (Aikhenvald 2000: 388). In the same way, Indonesian includes a natural gender marking system borrowed from Sanskrit, which can also be used with native stems (Stéphane Goyette, p.c.). Incidentally, gender agreement in adjectives referring to animates is precisely one of the morphological processes which DeGraff (2001a: 9–10, 2001b: 7–8) enumerates for Haitian.

5. The possibility of French influencing Haitian

As already mentioned, DeGraff is aware of the possibility of post-formative French influences, but he rejects it on the grounds that most speakers of Haitian are, and always have been, monolingual (DeGraff 2001a: 31, 2001c: 22).

---

13. Productivity, of course, is not a *sine qua non* in order for something to be an affix, but it does prove such a status.

This approach is convincing at first glance. However let us consider the case of another language, namely English. It is common knowledge that French, and to a lesser extent Latin and Greek, have had an impressive impact on English. Frequently cited figures concerning the non-inherited share of the English lexicon tend to be between 55% and 75% (e.g., Anttila 1972: 172). But then, of course, lexicon is more easily affected than the nucleus of a language – its core vocabulary and its grammar. The proportion of Romance items on the Swadesh 100-word list of English is 6%. Therefore, we might expect the Romance (and Greek) impact on English morphology to be of the same size. It is not. Naturally, it all depends on how you count, but applying the same criteria as DeGraff used when examining morphology in Haitian, it certainly isn't unreasonable to say that a majority of the productive affixes in English are of Romance, Greek, and – to a lesser extent – Scandinavian origin. Some occur only or mainly in more or less learned contexts, such as *ab-*, *ambi-*, *auto-*, *dys-*, *poly-*, *-ology*, *-vore* and *syn-*. Others, however, have become part of everyday speech, including *re-*, *-able*, *-tion*, *-er*, *-ist*, *-ity*, *-ous*, *dis-*, *-ic*, *-ify*, *-al*, *-in/-im/-irr* and *-itude*. In short, the number of borrowed affixes in English by far exceeds the total number of (allegedly inherited) affixes in Haitian.[14]

Now another question arises: if Haitians could not possibly have borrowed affixes from French because they do and did not in general speak French, then how many medieval Englishmen spoke French? Of course, no one knows for certain, but surely, any expert on English history would be surprised to learn that exposure to French in medieval Britain exceeded that in 18th to 20th century Haiti (Thomason & Kaufman 1988: 263–331; the wealth of references therein still provides the best overview of the influence of French on English).

Some readers might think that English is exceptional, and somehow a challenge to the Stammbaum theory. This, in my view, is mostly due to the Anglocentrism reigning in linguistics, as in all other sciences. It is not at all difficult to find European languages (let alone languages in the world as a whole) which show as strong signs of borrowings as English does.[15] As it happens, my own native tongue, Swedish, was influenced by

14. The 2002 edition of the *Oxford English Dictionary* lists 339 affixes, of which the overwhelming majority are borrowed from other languages. The borrowed affixes are derivational rather than inflectional, but so is the affix inventory of Haitian.

15. Limiting ourselves to the lexicon just to facilitate demonstration, other languages in which more than half of the total vocabulary is estimated to be borrowed include Armenian, Bai and Basque, while in Breton, Japanese, Korean, Maltese, Vietnamese and many Philippine languages, the proportion is over 60%. In Albanian, Brahui and Hungarian, the proportion is estimated at about 90% (Anttila 1972: 172; Brincat 2003; Dalby 1998: 36, 61, 668; Stéphane Goyette p.c.; Anthony Grant, p.c.; Hock 1991: 423; Matisoff 2001: 318; Shibatani 1990: 132–133; Sohn 1999: 87; Stolz 2002: 134; Trask 1996: 217, 309). These languages are by no means extreme – the same state of affairs obtains in more languages than most people are aware of.

Low German[16] in the Middle Ages to an extent perfectly comparable to that of English – about 6% on the Swadesh list, and around 60–70% of the lexicon as a whole (Anttila 1972: 172; Hock 1991: 423). Just as in English, it requires effort to maintain a conversation for more than a few sentences without using borrowed affixes. Again, I am confident that no one would claim that the number of Swedes with a knowledge of Low German in the Middle Ages ever exceeded the share of Haitians speaking French.

Other examples from medieval Europe include Latin influencing both German word order (Chirita 2003: 193), as well as providing 30% of the modern French lexicon (Posner 1996: 151, 1997: 164).[17] Even if we were to assume that the uneducated masses in England and Sweden were bilingual in the élite language to an extent that Haitians have never been – did the proletariat of Germany and renaissance France really speak Latin? Transfer of linguistic material does require bilinguals, but my point is that it does not require a large number of such people.

Linguists are well aware that while any language may borrow vocabulary from its neighbours, it takes a bit of pressure for a language to accept bound morphemes. Or rather, it takes more for this to happen. In fact, many languages have borrowed bound morphemes and/or morphological patterns.

A few examples from European languages which illustrate this are Welsh and Yiddish plural endings, the Basque participle, Estonian possessives and perfects, Germanic agentives, Lithuanian case suffixes, Serbian superlatives, several French derivational suffixes, Saami interrogatives, some of the Megleno-Romanian person endings and comparatives, a mood suffix, several derivations and possibly also a pluraliser in Mari (Anttila 1972: 170; Comrie 1981: 124, 1990: 11; Kangasmaa-Minn 1998: 244; Thomason & Kaufman 1988: 242–243; Trask 1996: 128, 310–311; Weinreich 1953: 31–32, 41–42). This list is by no means exhaustive.

Outside Europe, we might mention the Cantonese passive, Alsea pronoun affixes, the Amharic genitive, ergative markers in Alsea, Siuslaw and Coos, and the ergative alignment itself in Ngandi, the Tat agentive, conjunctions in Azerbaijani, Persian and Yupik, the Quechua plural suffix, gender agreement in Tagalog, Ilocano and Ndunga-le, most of the classifier system in Résigaro, as well as several affixes in Assamese, Persian, Afrikaans, Tajiki and Korean (Aikhenvald 2000: 383, 387; Campbell 1997: 246; Comrie 1981: 41, 84, 277; Edward Akhras, p.c.; Masica 1991: 212; Mithun 2000; Muysken 1986: 432; Posner 1996: 85; Sohn 1999: 217; Stolz 2002: 146–147, 149; Trask 1996: 311, 314; Weinreich 1953: 42; Wilkins 1996: 112; Windfuhr 1990: 114–115; Lekeberg 2002; Maarle 1996: 104, 113). Anyone familiar with the literature on language contact will be able to verify that most of the languages enumerated here are not the

16. The 'German' part of the name is partly misleading, as Low German has less in common with Standard German than it has with Dutch. One might even say that Dutch is the one dialect of Low German which happened to become standardised and elevated to the status of Language.

17. This refers to medieval Latin loans. Of course, inherited (as opposed to borrowed) Latin material forms the basis of the French word stock.

ones most frequently mentioned in discussions on heavy borrowing. Borrowing sometimes gets far more extreme than this. These are just a number of languages which have been in contact with others – something which any living language in the world is. The work criticised here focuses mainly on French and Haitian when it would be worthwhile to take other relevant cases into account. Note also that most of the cases mentioned concern inflectional rather than derivational morphology (which, being less tightly grammaticalised, is easier to borrow than inflections). This is of some significance, given that the Haitian affix inventory considered by DeGraff is exclusively derivational.

There is a certain ironic twist to this. In the papers already cited, DeGraff has emphasised that creoles are in no way unique, as the proponents of the broken transmission hypothesis suggest. However, if Haitian in fact did not undergo processes which have affected hundreds, if not thousands, of other languages, *that* would be unique.

DeGraff's view on the impossibility of Haitian having borrowed affixes from French is further complicated by the fact that one of those he himself enumerates, the agentive *-adò*, is, – as already mentioned – borrowed from Spanish (and DeGraff does indicate his awareness of this). Another agentive, *-mann* (for examples, see Valdman 1978: 137), which he does not mention, is clearly English in origin (deriving from the English noun *man*), and appears to have been borrowed from Jamaican Creole.

The population of Haiti, then, although massively monolingual, has obviously incorporated bound morphemes from the languages of the neighbouring countries (Spanish and English). In this light, it makes little sense to claim that French affixes cannot possibly have been borrowed into Haitian given that French is one of the country's official languages, while English and Spanish are not.

Continuity advocates may conceive of this as inconsistent with their interest in identity issues. Mufwene, for instance (interviewed in DeGraff 2001d), sees the very idea of decreolization as 'outrageous', 'pathological' and 'preposterous', since the "suggestion that people from the lower class aspire at speaking like those of the upper class is so contrary to sociolinguistic reality around us", and since creole speakers have "no social identity problem and have not wanted to be like [...] or speak a vernacular like [speakers of the prestige language]". I think this rather misses the point, since the incorporation of foreign linguistic material into one's own speech is not necessarily motivated by an urge to shift one's ethnic belonging. Rather, plenty of ordinary mortals are less reflective about identity issues, and simply pick up bits and pieces of surrounding speech habits in a not always conscious fashion. While speakers certainly do project sociolinguistic identities in their speech (as seen, for instance, in LePage and Tabouret-Keller's (1985) modern classic *Acts of Identity*) they do not live in a dichotomous universe where identities are mutually exclusive. Rather, plenty of modern work, from at least Labov (1966), illuminates how people create a fluid kind of identity from high to low with each utterance by exploiting resources from more than one register. Put differently, the idea that Haitians would 'want' to speak French is simplistic in view of modern sociolinguistic theory.

The Portuguese-lexicon creole of Annobón is an interesting case in point. The island was handed over to the Spanish in 1778, though there was no actual presence of Spanish speakers until the beginning of the 20th century. Spanish continues to be the official language of the island (since 1998 along with French). As a result, the Annobonese lexicon contains 10% Spanish elements (Bartens 1995: 126). Yet, there are no signs of language attrition – on the contrary, since the 18th century, the Annobonese have had a reputation of being fiercely independent, and even today, they take great pride in their language (Post 1994: 192; Lipski 1992: 37). Thus, the simple fact that the current prestige language on Annobón is different from the original lexifier (as opposed to the situation in Haiti) allows us to be certain that the Spanish component in it has entered the language after its formation. And yet, nothing suggests that the Annobonese are aiming at becoming a different people than they are, neither ethnically nor linguistically.

It might be added in this context that strong influence – including the borrowing of bound morphemes – from the prestige language into a creole has been attested for a number of creoles, including Guinea-Bissau Portuguese Creole, émigré Sranan, Papiamentu and Louisianais (Scantamburlo 1981: 39; Adamson 1998; Kouwenberg & Murray 1994: 5; Baum 1976: 83; Lipski 1986: 173; Grant 2002: 6; Morgan 1959; Neumann 1985; Neumann-Holzschuh 1987; Klingler 2000: 2). In addition, a copula has been demonstrated to have entered Mauritian French Creole from its lexifier more than a century after its creation (Baker & Syea 1991).[18] I propose that Haitian is not an exception to this process.

Finally, in the most well-known creoles which are spoken in isolation from their lexifier, the English-lexicon creoles of Suriname, in all about 700 English morphemes and not one single inherited affix have been attested (Smith 1987). Even if we assume that later borrowings from the subsequent prestige language, Dutch, ousted some English-derived material (for which there is preciously little evidence beyond a small number of lexical items), both the size of the inherited lexicon and the affix inventory are compatible with former pidginhood rather than with continuity. Given that both Suriname and Haiti were founded as plantation colonies, with rather similar socio-economic conditions (the main exception being the withdrawal of the lexifier in Suriname), the difference is intriguing. The same point could, once again, be made

18. More general references to decreolization are easily found for, e.g., the Portuguese-lexicon creoles of Asia and the Cape Verde Islands, the English-lexicon creoles of Hawaii, Trinidad and Belize, among many others (Hancock 1987: 280; Winer 1993: 12; Holm 1989: 286, 296–297, 524; Silva 1985). The phenomenon is of course not limited to European-lexicon creoles. It has been observed both in Sango (Christina Thornell, p.c.) and in Juba Arabic creole of southern Sudan, which Versteegh (1993) even describes as being on the way of becoming a dialect of Arabic. It is also worth noting that Lalla and D'Costa (1990: 37–46) show for Jamaican that although the language has become more basilectal in some areas (examples of this sort are often given with regard to Haitian by adherents of the continuity hypothesis), the trend in the *system as a whole* is clearly a drift towards the lexifier (in this case English).

with reference to Annobonese, with its single inherited affix. My suggestion is that it is not coincidental that the one language among these which remained in contact with its lexifier (i.e., Haitian) is also the one which shares the largest number of characteristics with it. If this suggestion is rejected, we are once more faced with the same question – why is Haitian different?

6. Conclusion

This paper has not proven that Haitian is a direct descendant of an expanded pidgin rather than of French. Nor do I claim to have shown that all (or even any) Haitian affixes are post-formative borrowings from French. It is probably impossible to produce watertight proof in favour of one or the other hypothesis. The best any side can do is to amass circumstantial evidence in favour of its position, and this is what I claim to have contributed to.

We have seen that a given Haitian dialect's degree of similarity to French in general correlates inversely with its similarity to the kind of French providing the superstratal input during the language's formation. We have also seen that agglutination, and in particular whole-syllable agglutination – a markedly non-French feature – is overrepresented in the basic lexicon rather than being evenly or randomly distributed across all nouns.

The data and analysis provided in this paper are highly suggestive of French having exercised a post-formative influence on Haitian, and this in turn makes present-day Haitian more compatible with the broken transmission hypothesis. The continuity hypothesis, on the other hand, would have more of a problem accounting for the observations made above. This applies in particular to the more extreme proponents of this position, who exclude the possibility of a latter-day French impact on Haitian. If French has indeed had a strong lexical influence on Haitian, and if Haitian behaves like other languages in contact do, this clearly allows the possibility of affixes having been borrowed as well. Again, this cannot be decisively proven, but at the very least it should be clear that such a possibility has not been convincingly tackled by the proponents of the continuity hypothesis. This is remarkable in view of the fact that the affix inventory of Haitian is probably the strongest linguistic (as opposed to socio-historical) evidence presented in favour of this position.

While it is true that certain socio-historical aspects of creole genesis in the Caribbean do lend themselves to the interpretation proposed by adherents of the continuity hypothesis, a discussion on the history of Haitian (or any other language) that gives precedence to socio-historical reconstructions over linguistic reality is bound to be incomplete at best. As linguists, we should rely first and foremost on languages, and only if the structure of a given language requires a particular explanation should we search for that in historical sources – history alone would never be able to, say, prove the Uralic identity of Hungarian or allow us to agree on the family status of Indo-European. It seems to me that if the map and the terrain fail to match, we should first

and foremost reconsider the map – in this case the reconstructed scenario of creole genesis.

An argument persisting in claiming that Haitian has by time become less and less like French must not only provide an account as to why Haitian would fail to behave in the way languages in contact usually do. It must also provide an alternative explanation to the distribution of syllable agglutination and to the fact that there is a significant inverse correlation between 'Frenchness' and 'archaicness'. The broken transmission hypothesis accounts for all these facts. That does not necessarily mean that it is correct, but for now, the burden of proof must lie on the alternative which leaves these issues unadressed.

Appendix

I first started out by tentatively defining the **Inner Core** as the most common words on the Haitian frequency list compiled by Jeff Allen and featured on the now defunct *Creolist Archives* web site. This was found inappropriate for two reasons. First, there was a bias towards religious texts, which led to plenty of proper names, often of Biblical characters and places, being featured. Secondly, items of the type featured on the Swadesh list, such as body parts, tend to be underrepresented in Bible texts and news stories (the other main source of Allen's data), even though no one would presumably deny that these are indeed basic. On the other hand, the Swadesh list omits most grammatical morphemes, including all bound ones. Therefore, the Inner Core here includes both the combined Swadesh 100- and 200-item lists (in all 207 items) and the top 500 of Allen's frequency list. In addition, the 500 most common words used in a 50 000 word corpus based on web-based discussions between Haitians was included. This material was chosen for want of a corpus of spoken Haitian, since it has been shown (e.g., McElhearn 1996; Baron 1998) that the genre in question tends to be closer to spoken language than most other types of writing.

For the **Outer Core**, five different corpora were examined, each containing about 50 000 tokens and 5 500 types. Any word occuring more than five times in three or more of these corpora, but not already featured in the Inner Core, was defined as belonging to the Outer Core. The original texts were identified through web searches including any group of five Inner Core items. To this were added the next 500 items on Allen's frequency list.

All in all, this process yielded an Inner Core lexicon of 789, and an Outer Core of 562 (these figures depending, of course, on how one defines a 'word', and how one deals with polysemy and other tricky aspects of word counting).

Finally, the **Periphery** was taken to include ranks 1 001–4 000 in the Allen list, along with the contents of Elias (2000) and an anonymously authored dictionary found at numerous web locations. The last source was used for the simple reason that I wanted word lists in electronic format rather than on paper, and the fact that inclusion in a dictionary guarantees that at least somebody has classified a given string as

a word rather than a misprint or a nonce formation. Those bothered by my use of an anonymous web source may note that this dictionary did indeed include several errors, which is important for two reasons. First, these errors appear to derive from the scanning of a printed (and thus, to many people by definition more reliable) source. Secondly, this neccessitated attentive proofreading with cross-checking in the printed dictionaries available to me. In practice, therefore, the Periphery consists of Haitian words as spelt and defined in standard dictionaries which are absent in the Core. The periphery amounts to 5 108 words.

In all the word lists, proper names not denoting specifically Haitian referents (typically words such as *Klintonn*) were excluded.

References

Adamson, L. (1998). Morphologische ontwikkelingen in het Sranan: Een gevolg van het contact met het Nederlands. In A. Bruyn & J. Arends (Eds.), *Mengelwerk voor Muysken* (pp. 282–286). Amsterdam: Institut voor Algemene Taalwetenschap, Universiteit van Amsterdam.

Aikhenvald, A. (2000). *Classifiers. A typology of noun categorization devices*. Oxford: OUP.

Alleyne, M. (1996). *Syntaxe historique créole*. Paris/Schœlcher: Éditions Karthala/Presses Universitaires Créoles.

Anon (1916). *Le français tel que le parlent nos tirailleurs sénégalais*. Paris: Imprimerie Militaire Universielle L. Fournier.

Anttila, R. (1972). *An Introduction to Historical and Comparative Linguistics*. New York NY: Macmillan.

Baker, P. & Corne, C. (1982). *Isle de France Creoles: Affinities and origins*. Ann Arbor MI: Karoma.

Baker, P. & Syea, A. (1991). On the copula in Mauritian Creole, past and present. In F. Bryne & T. Huebner (Eds.), *Development and Structures of Creole Languages* (pp. 159–175). Amsterdam: John Benjamins.

Baker, P. (1990). Off target? *Journal of Pidgin and Creole Languages, 5*, 107–119.

Baron, N. (1998). Writing in the age of email: The impact of ideology versus technology. *Visible Language, 32*(2), 35–53.

Bartens, A. (1995). *Die iberoromanisch-basierten Kreolsprachen*. Frankfurt: Peter Lang.

Baum, P. (1976). The Question of Decreolization in Papiamentu Phonology. *International Journal of the Sociology of Language, 7*, 83–93.

Brincat, J. (2003). Purism and neologisms in contemporary Maltese. In W. Boeder, J. Brincat, & T. Stolz (Eds.), *Purism in Minor Languages, Endangered Languages, Regional Languages, Mixed Languages* (pp. 55–170). Bochum: Brockmeyer.

Campbell, L. (1997). *American Indian languages: The historical linguistics of Native America*. Oxford: OUP.

Carayol, M. & Chaudenson, R. (1984/1989). *Atlas linguistiques de la France par régions: Atlas linguistique et ethnographique de la Réunion* (2 vols.). Paris: Éditions du Centre National de la Recherche Scientifique.

Chaudenson, R. (1979a). *Les créoles français*. Paris: Nathan.

Chaudenson, R. (1979b). Le français dans les îles de l'Océan Indien (Mascareignes et Seychelles). In A. Valdman & R. Chaudenson (Eds.), *Le français hors de France* (pp. 543–617). Paris: Éditions Honoré Champion.

Chaudenson, R. (1989). *Créoles et enseignement du français: Français, créolisation, créoles et français marginaux*. Paris: l'Harmattan.

Chaudenson, R. (1992). *Des îles, des hommes, des langues*. Paris: l'Harmattan.

Chirita, D. (2003). Did Latin influence German word order? Aspects of German-Latin bilingualism in the Late Middle Ages. In K. Braunmüller & G. Ferraresi (Eds.), *Aspects of Multilingualism in European Language History* (pp. 173–200). Amsterdam: John Benjamins.

Comrie, B. (1981). *The Languages of the Soviet Union*. Cambridge: CUP.

Comrie, B. (1990). Introduction. In B. Comrie (Ed.), *The Major Languages of Eastern Europe* (pp. 1–18). London: Routledge.

Comrie, B. (Ed.). (1990). *The Major Languages of South Asia, the Middle East, and Africa*. London: Routledge.

Corne, C. (1999). *From French to Creole*. London: University of Westminster Press.

Dalby, A. (1998). *Dictionary of languages*. London: Bloomsbury Publishing.

DeGraff, M. (2000). Creole speakers as a biologically definable class? Morphology in language creation. Paper presented at the Language in Society Workshop, University of Chicago, 10 January 2000. http://web.mit.edu/linguistics/www/degraff/uchicago.pdf. Accessed 2000-01-21.

DeGraff, M. (2001a). Morphology in creole genesis: Linguistics and ideology. In M. Kenstowicz (Ed.), *Ken Hale: A life in language* (pp. 53–121). Cambridge MA: The MIT Press. (Pagination here refers to the MS version at http://web.mit.edu/linguistics/www/degraff/festschrift.pdf, accessed 2002-11-03.)

DeGraff, M. (2001b). The abnormality of abnormal transmission in Creole genesis: Lessons from morpho-syntax and/in diachrony. Ohio State University, 23 May 2001. http://web.mit.edu/linguistics/www/degraff/lsa-handout-rev.pdf. Accessed 2002-11-03.

DeGraff, M. (2001c). 'Neo'-Darwinian Creolistics: A short debugging guide. Paper presented at the SPCL meeting, Coimbra, 27 June 2001.

DeGraff, M. (2001d). Salikoko Mufwene – from the Congo to Chicago. *Carrier Pidgin* 29.

DeGraff, M. (2002). Relexification: A reevaluation. *Anthropological Linguistics, 44*, 321–414.

Dorrance, W. A. (1935). *The Survival of French in the Old District of Ste-Geneviève*. Columbia: University of Missouri.

Elias, W. L. (2000). Ayiti an kreyòl/diksyonè a. http://ayitionline.freeservers.com/Alfabet.htm. Last accessed 2004-03-10.

Fattier, D. (1998). *Contribution à l'étude de la genèse d'un créole: l'atlas linguistique d'Haïti, cartes et commentaires*. Villeneuve d'Asq: Presses universitaires du Septentrion.

Goddard, I. (1996). Pidgin Delaware. In S. Thomason (Ed.), *Contact Languages. A wider perspective* (pp. 43–98). Amsterdam: John Benjamins.

Grant, A. (1995). Article agglutination in Creole French: A wider perspective. In P. Baker (Ed.), *From Contact to Creole and Beyond* (pp. 149–176). London: University of Westminster Press.

Grant, A. (1996). The evolution of functional categories in Grand Ronde Chinook Jargon: Ethnolinguistic and grammatical considerations. In P. Baker & A. Syea (Eds.), *Changing Meanings, Changing Functions* (pp. 225–242). London: University of Westminster Press.

Grant, A. (2002). A Constructivist Approach to the Early History of Papiamentu. Ms.

Griolet, P. (1986). *Mots de Louisiane – Étude lexicale d'une francophonie*. Gothenburg/ Paris: Acta Universitatis Gotoburgensis/ l'Harmattan.

Hall, R. (1966). *Pidgin and Creole Languages*. Ithaca NY: Cornell University Press.

Hancock, I. (1987). A preliminary classification of the Anglophone Atlantic Creoles, with syntactic data from thirty-three representative dialects. In G. Gilbert (Ed.), *Pidgin and Creole Languages* (pp. 264–333). Honolulu HI: University of Hawaii Press.

Highfield, A. (1979). *The French Dialect of St. Thomas, U.S. Virgin Islands. A descriptive grammar with texts and glossary*. Ann Arbor MI: Karoma.
Hock, H. H. (1991). *Principles of Historical Linguistics*. (2nd revised and updated edition). Berlin: Mouton de Gruyter.
Hollyman, J. (1964). *Les français régionaux de l'Indopacifique: Essais de phonologie.* Auckland: Linguistic Society of New Zealand.
Holm, J. (1989). *Pidgins and Creoles* (Vol. 2.). Cambridge: CUP.
Kangasmaa-Minn, E. (1998). Mari. In D. Abondolo (Ed.), *The Uralic Languages* (pp. 219–248). London: Routledge.
Kaye, A. & Tosco, M. (1993). Early East African Pidgin Arabic. *Sprache und Geschichte in Afrika, 14*, 269–305.
Klingler, T. (2000). Louisiana Creole: The multiple-genesis hypothesis reconsidered. *Journal of Pidgin and Creole Languages, 15*, 1–35.
Kouwenberg, S. & Murray, E. (1994). *Papiamentu*. Munich: Lincom.
Kouwenberg, S. (Ed.) (2003). *Twice as Meaningful*. London: Battlebridge.
Labov, W. (1966). *The Social Stratification of English in New York City*. Washington DC: Center for Applied Linguistics.
Lalla, B. & d'Costa, J. (1990). *Language in Exile. Three hundred years of Jamaican Creole*. Tuscaloosa AL: University of Alabama Press.
Lefebvre, C. (1998). *Creole Genesis and the Acquisition of Grammar*. Cambridge: CUP.
Lekeberg, Z. (2002). On grammatical diversity in China: Passive constructions in Sinitic languages. MS, Stockholm University.
LePage, R. & Tabouret-Keller, A. (1985). *Acts of Identity: Creole-based approaches to language and ethnicity*. Cambridge: CUP.
Lipski, J. (1986). Convergence and divergence in Bozal Spanish. *Journal of Pidgin and Creole Languages, 1*, 171–203.
Lipski, J. (1992). Pidgin English usage in Equatorial Guinea (Fernando Poo). *English World-Wide, 13*, 33–57.
Maarle, J. van (1996). On the interplay of inherited and non-inherited features in Afrikaans derivational morphology. In H. Nielsen & L. Schøsler (Eds.), *The Origins and Development of Emigrant languages* (pp. 103–115). Odense: Odense University Press.
Masica, C. (1991). *The Indo-Aryan languages*. Cambridge: CUP.
Massignon, G. (1962). *Les parlers français d'Acadie: Enquête linguistique*. Paris: Klincksieck.
Matisoff, J. (2001). Genetic versus contact relationship: Prosodic diffusibility in Southeast Asian languages. In A. Aikhenvald & R. Dixon (Eds.), *Areal Diffusion and Genetic Inheritance* (pp. 291–327). Oxford: OUP.
McElhearn, K. (1996). Writing Conversation: An analysis of speech events in e-mail mailing lists. MA thesis, Aston University.
Mithun, M. (2000). Ergativity and language contact in Oregon: Alsea, Siuslaw, and Coos. Paper presented at Stockholm University, 26 May 2000.
Morgan, R. (1959). Structural sketch of St. Martin Creole. *Anthropological Linguistics, 1*, 8–20.
Mufwene, S. (1992). Africanisms in Gullah: A re-examination of the issues. In J. Hall, N. Doane, & R. Ringler (Eds.), *Old English and New: Essays in language and linguistics in honor of Frederic G. Cassidy* (pp. 156–182). New York NY: Garland.
Muysken, P. (1996). Callahuaya. In S. Thomason (Ed.), *Contact Languages. A wider perspective* (pp. 427–447). Amsterdam: John Benjamins.
Neumann, I. (1985). *Le créole de Breaux Bridge, Louisiane*. Hamburg: Helmut Buske.

Neumann-Holzschuh, I. (Ed.). (1987). *Textes anciens en créole louisianais*. Hamburg: Helmut Buske.

Parkvall, M. (2001). Interlanguage no be pidgin. Paper presented at the Workshop on Pidgin languages – their nature and significance, London, April 2001.

Parkvall, M. (2003). Reduplication in the Atlantic creoles and other contact languages. In S. Kouwenberg (Ed.), *Twice as Meaningful* (pp. 19–36). London: Battlebridge.

Posner, R. (1996). *The Romance Languages*. Cambridge: CUP.

Posner, R. (1997). *Linguistic Change in French*. Oxford: Clarendon Press.

Post, M. (1994). Fa d'Ambu. In J. Arends, P. Muysken, & N. Smith (Eds.), *Pidgins and Creoles. An introduction* (pp. 191–204). Amsterdam: John Benjamins.

Prokosch, E. (1986). *Arabische Kontaktsprachen (Pidgin- und Kreolsprachen) in Afrika*. Graz: Universität Graz.

Scantamburlo, L. (1981). *Gramática e dicionário da língua criol da Guiné-Bissau* (*GCr*). Bologna: Editrice Missionaria Italiana.

Shibatani, M. (1990). Japanese. In B. Comrie (Ed.), *The Major Languages of South Asia, the Middle East, and Africa* (pp. 127–152). London: Routledge.

Siegel, J. (1987). *Language Contact in a Plantation Environment*. Cambridge: CUP.

Silva, I. S. (1985). *Variation and Change in the Verbal System of Capeverdean Crioulo*. Ph.D dissertation, Georgetown University.

Smith, N. (1987). The Genesis of the Creole Languages of Suriname. Ph.D dissertation, Universiteit van Amsterdam.

Société du parler français au Canada (1930 [1968]). *Glossaire du parler français au Canada*. Québec: L'Action Sociale/Presses de l'Université Laval.

Sohn, H.-M. (1999). *The Korean Language*. Cambridge: CUP.

Stolz, T. (2002). General linguistic aspects of Spanish-indigenous language contacts with special focus on Austronesia. *Bulletin of Hispanic Studies, 79*, 133–158.

Thogmartin, C. O. (1979). Old Mines, Missouri, et la survivance du français dans la haute vallé du Mississippi. In A. Valdman & R. Chaudenson (Eds.), *Le français hors de France* (pp. 111–118). Paris: Éditions Honoré Champion.

Thomason, S. & Kaufman, T. (1988). *Language Contact, Creolization, and Genetic Linguistics*. Berkeley CA: University of California Press.

Thomason, S. (1980). On Interpreting 'The Indian Interpreter'. *Language in Society, 9*, 167–193.

Trask, L. (1996). *Historical Linguistics*. London: Arnold.

Valdman, A. & Chaudenson, R. (Eds.). (1979). *Le français hors de France*. Paris: Éditions Honoré Champion.

Valdman, A. (1978). *Le créole: Structure, statut et origine*. Paris: Klincksieck.

Van der Voort, H. (1997). New light on Eskimo pidgins. In A. Spears & D. Winford (Eds.), *The Structure and Status of Pidgins and Creoles* (pp. 373–394). Amsterdam: John Benjamins.

Weinreich, U. (1953). *Languages in Contact. Findings and problems*. New York NY: Linguistic Circle of New York.

Versteegh, K. (1993). Leveling in the Sudan: From Arabic Creole to Arabic dialect. *International Journal of the Sociology of Language, 99*, 65–79.

Wilkins, D. (1996). Morphology. In H. Goebl, P. Nelde, Z. Starý, & W. Wölk (Eds.), *Kontaktlinguistik. Ein internationales Handbuch zeitgenössischer Forschung* (pp. 109–117). Berlin: Walter de Gruyter.

Windfuhr, G. (1990). Persian. In B. Comrie (Ed.), *The Major Languages of South Asia, the Middle East, and Africa* (pp. 108–131). London: Routledge.

Winer, L. (1993). *Trinidad and Tobago* [Varieties of Engish around the World T6]. Amsterdam: John Benjamins.

Vitale, A. (1980). Kisetla: Linguistic and sociolinguistic aspects of a pidgin Swahili of Kenya. *Anthropological Linguistics, 22*, 47–65.

Zengel, M. S. (1962). Literacy as a factor in language change. *American Anthropologist, 64*, 132–139.

The late transfer of serial verb constructions as stylistic variants in Saramaccan creole

Marvin Kramer
Dharma Realm Buddhist University, Ukiah, CA, USA

Verb serialization in Saramaccan is a transferred feature from Fongbe, where in both languages there are non-serial alternatives to serial verb constructions (SVCs). McWhorter (2002) argues that only features minimally necessary to language would transfer during creole genesis, isomorphically precluding stylistic variants. In a pidgin stage with speakers of English and Fongbe there would be no need for SVCs, as both languages have the same basic word order. But in modern Saramaccan and the historical texts there are non-serial constructions with SVCs as stylistic variants, not an unexpected aspect of creole expansion. Verb serialization would then have transferred during creole expansion, when the language was primarily a medium for community solidarity by Africans, and would represent the transfer of a non-essential feature.

1. Introduction

McWhorter (2002) argues for a model of creole genesis in which the creole's earlier pidgin phase acts as a filter on expansion by blocking the transfer of some features. In this view, these features are not available to emerge during creole genesis. The features thus blocked are precisely those that develop over time, and are therefore deemed 'ornamental' (p. 4), hence 'nonessential' to language (34). Serial verb constructions (SVCs) are not included among the specific features discussed, but SVCs ought also to be excluded by the pidgin filter as they do not fit into McWhorter's list of features considered necessary in a natural language (7). Assuming isomorphism to be a pidgin trait, the list specifies single features, like 'a subordination strategy' or 'a focus marking strategy', but SVCs in Saramaccan have non-serial alternative clauses that would render SVCs nonessential. SVCs transferred into the Suriname creoles from the Kwa languages, the most significant substrate languages being the Gbe languages (Arends 1995), and in particular Fongbe (Migge 1998). Hyman (2000: 15) argues that SVCs in Kwa developed over time as a result of the loss of inflectional morphology. The SVCs that would transfer into the Suriname creoles, then, would be considered nonessential. If, however, SVCs had transferred subsequent to the pidgin phase, the pidgin filter

could still be seen as effective. In this case, the existence of SVCs in Saramaccan would not present a counterexample to the pidgin filter. McWhorter argues (p. 33), however, that creole genesis would not involve substrate transfer of nonessential features after the application of the pidgin filter, and so this model of creole genesis does not account for the existence of SVCs in Saramaccan.

Arguments proposing the existence of a common pidgin origin of the creoles in the English Caribbean, e.g. Caribbean Plantation Pidgin (Smith 2002: 134), cite the many similarities unique to the Caribbean English Creoles. In addition, McWhorter (2002) also argues for a pidgin stage based on the absence of 'ornamental' features in the subsequent creoles. The pidgin stage may filter out features that are not common or intelligible to the creators of the pidgin (Thomason & Kaufman 1988: 192). As a medium for interethnic communication (MIC; Baker 2000), the pidgin predecessor to Saramaccan would lack SVCs because they would not be familiar to speakers of European languages, or to speakers of other substrate non-serializing African languages like Kikongo, while non-serial alternatives would be familiar and intelligible to speakers of Kikongo and the European languages as well as the Kwa languages. Speakers of non-serializing languages would not be inclined to respond with SVCs, so they would not be reinforced. But as the pidgin expanded to become a medium for community solidarity (MCS) (Baker 2000), i.e. a language whose principal function is the expression of community solidarity (ibid.: 48) for a population that would by then have come to have a majority of Kwa speakers (Arends 1995: 243) and would naturally exclude Europeans, SVCs would be familiar and intelligible, and would provide stylistic alternatives. The MCS probably began to develop during the period of English control of Suriname, between 1651 and 1680, when the ratio of Africans to Europeans was small and communication between them could have been in an L2 English by a diglossic slave population (McWhorter 2000: 99–104) whose MCS would exclude Europeans. The date of the eventual withdrawal of most of the British from Suriname, around 1680 (Arends 2002: 119–121), would then mark the end of English influence on any MIC but not necessarily the end of the MCS stage of the creolization process. Many British returned to Suriname after 1680, but by then it had become a multilingual society with no single language of power (Arends 2002). Furthermore, at this time there was a major increase in the African population (Arends 1995: 243), greatly reducing any lexifier influence, and thereby facilitating the transfer of SVCs. McWhorter (p. 33) argues that the pidgin filter applies throughout creolization, but a transfer of SVCs in the MCS stage assumes that creolization can involve substrate transfer of nonessential features after the pidgin filter and during the period of creole expansion.

An indication that the pidgin stage had no SVCs is the existence of non-serial alternatives for all types of SVCs in Saramaccan, either in the modern language or in the early texts. These SVCs and their non-serial alternatives are described in detail in Section (2). There are also non-serial alternatives for these SVCs in Fongbe (Lefebvre & Brousseau 2002), the major substrate language of the Surinamese English creoles (Migge 1998). Because the Gbe languages are SVO their non-serial constructions are SVO constructions, and when transferred would be familiar and intelligible to both

English and Fongbe speakers in an MIC pidgin, and would suffice for limited MIC functions. Input from the socially dominant speakers of English in particular would require no compromise on the basic SVO word order.

An expanded MCS, without further input from English, would incorporate substrate SVCs. This transfer of SVCs from Fongbe into the creole predecessor to Saramaccan, however, resulted in SVCs in Saramaccan that are either as grammaticalized as the corresponding SVCs in Fongbe, or are less grammaticalized. As will be seen in Section (2), in the process of transfer particular SVCs have undergone what in normal diachrony would appear to be the reverse direction of grammaticalization, in that the verbs in the SVCs in Saramaccan have more verbal characteristics than the verbs in the corresponding SVCs in Fongbe. Bruyn (1996) demonstrates that what would appear to be grammaticalization of complex PPs in the related Surinamese creole Sranan is in fact the transfer of grammaticalized forms. In Saramaccan, however, the transfer of SVCs has not included the transfer of some of the grammaticalization of those SVCs. For a number of SVCs in Fongbe in particular, the loss of the verbal property of clefting for V2s, the loss of verbal properties for prepositional properties for V2s, and the lexicalization of serial verbs into single semantic representations did not transfer into the corresponding SVCs in Saramaccan. The increase in verbal properties and loss of lexicalizations in the transfer of SVCs into Saramaccan could be explained not so much as by the application of a pidgin filter as such, but rather by the fact that speakers of many serializing languages as well as non-serializing languages, such as Kikongo, were part of the speech community using the MCS.

McWhorter (2001, 2004) argues, using Saramaccan as an example, that the failure of serial verb lexicalization to be represented in transferred SVCs is a creole characteristic, namely simplification. The semantics of verbs in creole SVCs have not fused beyond accessibility, in contrast to many SVCs in the substrate, so that the verbs in creole SVCs more clearly represent 'sub-events' (Durie 1997). Semantically, these SVC verbs are more compositional, and thus closer to their corresponding use in non-serial depictions than lexicalized SVC verbs. The semantic transparency of verbs in creole SVCs is perhaps due to the fact that speakers of many different languages were involved, particularly speakers of languages that were not primary substrate languages for whom the attempted use of lexicalizations would be an impediment to communication, the 'tertiary hybridization' of Whinnom (1971: 105–106). In other words, although SVCs were common to the Gbe languages, particular lexicalizations of verbs in these SVCs were not. The absence of lexicalized verbs in SVCs in creole languages allows a clearer equivalence with those verbs in non-serial clauses, where there can be variation, than with lexicalized SVC verbs.

Durie's typology of SVCs highlights a number of similarities between SVCs and non-serial clauses (1997: 291). Both serial and non-serial clauses depict what is loosely defined as a single event; both have a single tense, aspect, modality and polarity; both have the same intonational properties (Givón 1991). Both have only one external and one internal argument. And when there are more than two arguments, the configuration of arguments in serializing languages corresponds to the configuration of

arguments and adjuncts in non-serializing languages. SVCs, presumably those with lexicalized verbs, can often be best translated by single verb clauses in non-serializing language. But Durie (p. 328) points out that there are sequences of sub-events that are culturally and thereby linguistically accepted as single events, and there are others that are not. The latter do not occur in SVCs, but are depicted in sequences of single verb clauses. There are, of course, linguistically unacceptable sequences for SVCs in Saramaccan as well (Kramer 2002: 225–235, 548–551; and in (3) below). Often, however, a minor alteration, such as the addition of an adverb (Durie, p. 329) or a different interpretation of the event (ibid.: 550), allows an otherwise unacceptable sequence of verbs to be acceptable in an SVC. Subtle interpretations of linguistically acceptable sequences of sub-events, then, can determine any variation between an SVC with non-lexicalized verbs and a sequence of single verb clauses with those verbs.

Considering that SVCs in creole languages have only non-lexicalized verbs, the variation between SVC and clause sequences would apply to most potential SVC depictions. There is also the possibility of variation between an SVC and a single verb clause, where the unexpressed verb is understood. In both the creole Bislama and the substrate Paamese (Crowley 2002: 241–242), for example, a clause with the verb *send* may occur without the verb *go*, although the SVC with both verbs is stylistically preferred. This stylistic preference for an SVC with *send* and *go* is exactly parallel to Saramaccan, where the SVC is preferred but a single verb clause with *send* is a stylistic variant, as in (7d) below. To the degree that creole SVCs do not have lexicalized verbs, then, there should be non-serial alternatives for SVCs.

There is undoubtedly a difference between SVCs and non-serial alternatives, such as, for example, Durie's (1997) observation that SVCs signal a culturally and linguistically acceptable sequence of sub-events, or Crowley's observation that SVCs can be stylistically preferred. Lambrecht (1995: 6) refers to such variations as 'allosentences', where there is more than one way to say basically the same thing. The occurrence of allosentences would be a departure from the isomorphic precept of the pidgin stage, namely that for any given meaning there is only one expression. McWhorter (1996) argues that SVCs "do not constitute a grammatical function *per se:* they are simply a grammar-wide configurational strategy which in turn serves to encode a range of functions" (p. 155). But if a non-serial clause can handle essential functions, then a second and arguably more marked syntactic device would be nonessential.

According to McWorther (2002), the determination that any feature is nonessential or not fundamental to grammar is not based on its function, but on its cross-linguistic occurrence. Some features, like grammatical gender (p. 4), carry little functional load, while others, like markers of spatial location (p. 5), convey very definite meanings. Both, however, are considered nonessential to language because they do not occur in every language. Even tense and aspect marking are not considered essential because they do not occur in all languages (p. 8). These, then, are features that develop over time and show up only in particular languages. SVCs similarly convey meaning that non-serial alternatives do not, and do not occur in every language, and, at least for the Kwa languages, have developed over time (Hyman 2000).

One of the bases for considering SVCs nonessential would be their use as stylistic alternatives, since their occurrence would be optional. The nonessential features discussed in McWhorter (2002), however, are not optional, it is argued, and so these features are not in creoles not because they occur too infrequently to be acquired, but rather because they are nonessential (p. 3). However, some of these features are optional in some languages. The examples discussed look at analytic marking, so lack of transfer would not be due to loss of inflection. Regarding the discussion of Fongbe alienable/inalienable possession (14), both are marked analytically either by the genitive marker *tɔ̀n*, or by the defective paradigm of POSS pronouns, which includes only the 1st and 2nd person sg.; alienable possession may optionally be marked by the Object marker *sín* (Lefebvre & Brousseau 2002: 71). Interestingly, there is also alienable/inalienable possession in Saramaccan, argued to be transfer by Kramer (2006), or to be an internal development by McWhorter (2004: 99). In Saramaccan, both alienable and inalienable possession use the preposition *fu* ~ *u* 'for/of', in example (1), but inalienable possession may optionally use juxtaposed POSS pronouns or NPs. Nouns that are inalienably possessed include items like clothing and houses as well as body parts and kinship terms, not unusual in the typology of inalienable possession (Nichols 1992: 120):

(1)	a.	i.	dí wósu u mí the house of me 'my house'	ii.	dí u mí wósu the of me house 'my house'	iii. mí wósu my house 'my house'
	b.	i.	dí búku u mí the book of me 'my book'	ii.	dí u mí búku the of me book 'my book'	iii. *mí búku my book (my book)

Similarly, the evidential marker in Hixkaryana (McWhorter 2002: 20) is optional (2). In 3rd person narrative in creation and animal stories where it would be most expected, for example, it is commonly absent, as in stories about creation (Derbyshire 1979: 47), about sloth (p. 59), jaguar chiefs (p. 74), and buzzard people (p. 77):

(2) txorofofo yk̨ihtoy ha karaw xarha
(black-nunbird he-made-it intensfr. (species of) bird add.)
'He created the black nunbird and the (species of) bird.' (Derbyshire p. 47)

Another listed feature is ergativity. McWhorter (2002) has argued, using Tongan as an example, that ergativity is not a syntactic option (p. 9). But there is an example in the closely related language Samoan of the pragmatically motivated variation in the use of ergative and genitive marking (Duranti 1990). As for frequency of occurrence, ergative marking is generally rare due to the discourse dispreference for introducing agents (Du Bois 1987), and the overwhelming tendency for ergative languages to have accusative-aligned pronominal paradigms. Their use of ergative marking, then, would be a marked stylistic option involving the naming of agent arguments. For any feature that develops over time, there must be an initial period when its use is an optional stylistic variation, along the notion that "grammar is style" (Labov 1990: 45). And if

anything, it is the optional stylistic variation that could be considered ornamental and nonessential. If SVCs in Saramaccan are a stylistic option, it stands they would not have passed the pidgin filter. Their presence in Saramaccan would then mean transfer after the hypothetical pidgin stage, or in any case during the expansion of the MCS stage.

Another indication that SVCs transferred relatively late, namely after the pidgin stage but before the late 18th century, is their apparent increase in frequency since the early texts of the late 18th century, discussed in Section 2 below. The incorporation of a feature, by transfer or grammaticalization, would be followed by an increase in the frequency of that feature; for example, since the 18th century the Sranan demonstrative *da* has increased in frequency (Bruyn 1996: 39), a process that most likely began after a pidgin stage, in that the lack of a definite article is characteristic of early pidgins (McWhorter 1996: 162). On the other hand, there has been no such increase in the frequency of the indefinite singular marker *wan*, indicating that its development from the numeral 'one' must have occurred early in Sranan (Bruyn 1996), that is, before the development of *da*. The increase in frequency of SVCs in Saramaccan should indicate a similar late incorporation. Of course the frequency of different SVCs has increased at different rates, in Section 2, suggesting incorporation at different times. The goal or beneficiary (GOAL/BEN) SVC with V2 *dá* 'give', for example, in 2.2.2, has completely replaced the non-serial alternative with the preposition *na* since the late 18th century, suggesting that it entered the language earlier than other SVCs. The rate of increase for the SVC with V2 *dá* could also have been affected by the fact that this SVC, unlike others, transferred with a partially grammaticalized V2, as *takki da* in the early texts means 'talk to', not 'talk give'. This head start on grammaticalization would reduce the semantic difference with the non-serial alternative with *na* 'at / to', leading to its eventual replacement by eliminating the option for contrast with the SVC.

The data on modern Saramaccan comes from Kramer (2002, 2005), Aboikoni and Glock (1997), Veenstra (1996), Rountree (1972), the Summer Institute of Linguistics (SIL) Saramaccan translation of the New Testament (1991), and my own unpublished field notes. Historical data is from the Wietz' 1805 Saramaccan translation of the Book of Acts published in Schuchardt (1914), the 1778 Schumann Saramaccan Dictionary (also in Schuchardt), the c. 1790 Riemer Saramaccan Dictionary (published in Arends & Perl 1995), and the 1790–1791 Maroon Letters (also in Arends & Perl).

2. SVCs and corresponding non-serial alternatives

In Sections 2.1 and 2.2 examples will be given to argue that there are non-serial alternative constructions for all the types of SVCs in Saramaccan, either in the modern language or in the historical texts.

The preference for the use of serial or non-serial constructions varies by the type of SVC. Based on informants' intuition and preferences, the comparison SVC, for example, is relatively rare, while for the directional SVCs it is the non-serial alternatives

that are rare. The GOAL/BEN SVCs with V2 *dá* 'give' have no non-serial alternative in modern Saramaccan, but in the early texts are apparently slightly less preferred than the non-serial alternative with the preposition *na* 'at/to' – with 29 occurrences of *da* 'give' with *takki* 'talk', compared to 33 with *na* 'on, at' – in the Wietz (1805) text, as discussed in 2.2.2 below. For each set of SVC and non-serial alternative, it is the SVC that is restricted in some way where the non-serial alternative is not. The most basic restrictions on SVCs, of course, are that they must depict culturally acceptable sequences of sub-events, and that they must depict single events, however loosely that is interpreted. The non-serial alternatives, then, may occur where the SVCs could occur, and elsewhere.

The constraints on SVCs in Saramaccan parallel constraints in Fongbe, suggesting that the constraints transferred rather than developed over time through grammaticalization in the creole language. From the outset of their occurrence, then, the use of SVCs would be limited compared to non-serial alternatives. Some of the conditions that limit the use of SVCs will be discussed in Sections 2.1 and 2.2, but the important point is that there is variation between SVCs and non-serial alternatives, and that this variation means the more limited option is not essential.

2.1 SVCs in Modern Saramaccan

2.1.1 *Comparison SVCs*

The standard of comparison for the comparative SVC is a VP, so that the verb *pasá* 'pass' is a V2 in a SVC; in the examples in (3) the property items are intransitive stative verbs. In (3) and subsequent examples the SVC is (i) and the non-serial alternative is (ii):

(3) a. i. a bígi pasá mí
3sg big pass me
'he is bigger than me'
ii. a bígi mɔ̃́ɔ̃ mí
3sg big more me
'he is bigger than me'
b. i. *a pikí pasá mí
3sg small pass me
(he is smaller than me)
ii. a pikí mɔ̃́ɔ̃ mí
3sg small more me
'he is smaller than me'
c. i. a kɔ́tɔ pói
3sg cold spoil
'it's too cold'

ii. a kɔ́tɔ túmísi
3sg cold too much
'it's too cold'

The SM V2 *pasá* is limited by its meaning to depictions suggesting growth or movement, thus excluding *pikí* in (3b.i) (Kramer 2002: 194). The alternative with *mɔ̃́ɔ̃* is not so constrained. The Fongbe V2 *hú* 'surpass' is more grammaticalized than Saramaccan *pasá*, allowing *hwè hú* 'short surpass', 'shorter' (Lefebvre & Brousseau 2002: 435). Similarly, the VP2 in the Fongbe comparative SVC is more grammaticalized than simple SVCs in having syntactic properties of PPs (p. 436), and the Fongbe superlative SVC with V2 *dín* has syntactic properties of PPs (p. 438) as well, allowing the serial and non-serial interpretations in (3). In addition, the Fongbe comparative serials have a more grammaticalized V2 than Saramaccan in that, unlike Saramaccan, it cannot be predicate clefted (p. 437). In the transfer of the comparison SVC, then, the V2 in Saramaccan has more syntactic and semantic verbal characteristics than the corresponding V2 in Fongbe, as predicted in McWhorter (2004), which, in a normal diachronic context from Fongbe to Saramaccan, would appear to be the reverse direction of grammaticalization.

In Saramaccan, *mɔ̃́ɔ̃* can be a full verb meaning 'to exceed', but the alternative constructions with *mɔ̃́ɔ̃*, with an NP2 standard of comparison as in (3a.ii) and (3b.ii), are not SVCs, as determined by tone sandhi; verbs sandhi with the following pronominal object, but *mɔ̃́ɔ̃* does not (Rountree 1972: 322); and by the avoidance of stranding, a characteristic of prepositions in Saramaccan. The non-serial PP characteristics of the Fongbe comparative, as well as the use of *mɔ̃́ɔ̃* from the etymon *more*, would allow a non-serial interpretation by speakers of both English and Fongbe, without the semantic constraints of the SVC with *pasá*, considering, of course, that the *mɔ̃́ɔ̃* construction was not an SVC in an earlier time.

In the variation between the *pasá* SVC and non-serial alternatives, it would appear that in the early texts there was little preference for the SVC. It does not occur in the 1805 Wietz Bible (Acts 1–13:33), the 1778 Schumann Dictionary, the Riemer Dictionary *c.* 1790, or the Maroon letters *c.* 1791, or in fact in the early texts of the closely related Sranan. This lack of preference in the late 18th century could indicate the comparison SVC with *pasá* would be a good candidate for late transfer into Saramaccan. Arends (1989: 85) finds that in Sranan the corresponding SVC with *psa* did not occur until the beginning of the 20th century, relexifying *moro* in some functions. The *pasá* SVC could similarly have entered Saramaccan at a late date, but it would not necessarily be a relexification of *mɔ̃́ɔ̃*, as the *mɔ̃́ɔ̃* construction may not have been an SVC. In fact, there is no occurrence of the *mɔ̃́ɔ̃* (*morro*) construction in the early Saramaccan sources, but occurrences in early Sranan suggest that it may have existed in early Saramaccan as well; *a biggi morro mi* 'he's fatter than me' (Schumann 1783, cited in Kramp 1983: 62). Comparisons in the early Saramaccan texts use the non-verbal *morro* 'more', *morro kumma* 'more like', or *morro leki* 'more like', even though in the Riemer and Schumann Saramaccan dictionaries *morro* can be a full verb, as in modern

Saramaccan. There are also other non-serial comparative constructions in early Saramaccan, such as using the preposition *vo* 'of, for'; *ju grandi vo mi* 'you're older than me' (Schuchardt 1914: 67). Similarly with the SVC with *pói* 'spoil', in (3c) above, the early texts use variations of *tumushi* 'too much' rather than an SVC with *poli* 'spoil'. It would seem, then, that the SVCs of comparison did not occur commonly until quite late, suggesting late transfer.

2.1.2 *Aspectual 'Finish' SVCs*

The aspectual SVC with V2 *kabá* 'finish' has a non-serial alternative construction with *kabá* as the matrix verb, which, as with other non-serial alternatives, is not constrained like the SVC. In (4b.i) and (4c.i), for example, the *kabá* 'finish' SVC does not occur with non-specific predicates. However, the SVC, by speaker intuition, may imply completion, while the *kabá u* construction may imply an intended action not completed. Other non-serial alternatives are clauses with *kàà* 'already' or variations on *té kabá* 'until finished':

(4) a. i. a fố dí alísi kabá
3sg pound the rice finish
'she finished pounding the rice'

ii. a kabá u fố dí alísi
3sg finish for pound the rice
'she finished pounding the rice'

iii. a fố dí alísi kàà / té a kabá
3sg pound the rice already / until 3sg finish
'she pounded the rice already / until she finished'

b. i. *a fố alísi kabá
3sg pound rice finish
(she finished pounding rice)

ii. a kabá u fố alísi
3sg finish for pound rice
'she finished pounding rice'

iii. a fố alísi kàà / té a kabá
3sg pound rice already / until 3sg finish
'she pounded rice already / until she finished'

c. i. *mi mbéi a wóóko kabá
1sg make 3sg work finish
(I made him finish working)

ii. mi mbéi a kabá u wóóko
1sg make 3sg finish for work
'I made him finish working'

iii. mi mbéi a wóóko té a kabá
1sg make 3sg work until 3sg finish
'I made him work until he finished'

There are corresponding serial and non-serial Fongbe aspectual 'finish' constructions (Lefebvre & Brousseau 2002: 434). SVCs in Fongbe and Saramaccan are sensitive to aspectual classes based on semantic properties (85), described by Dowty (1979) and others as STATES, ACTIVITIES, ACHIEVEMENTS and ACCOMPLISHMENTS; STATES (*know, be sick*) and ACTIVITIES (*eat pizza*) have no natural endpoint, ACHIEVEMENTS (*catch sight of*) depict a change of STATE as the endpoint result of action, and ACCOMPLISHMENTS (*eat the pizza*) depict the action and the change of STATE. The Saramaccan 'finish' SVC depicts only ACHIEVEMENTS and ACCOMPLISHMENTS. It does not depict ACTIVITIES, in (4b.i) and (4c.i), or STATES (Byrne 1987: 224), and speakers accept ACHIEVEMENTS reluctantly. The non-serial alternative to the *kabá* SVC in Saramaccan is not so constrained.

The Fongbe aspectual 'finish' SVC allows ACTIVITY and ACCOMPLISHMENT depictions only (Lefebvre & Brousseau 2002). By allowing ACTIVITY depictions the Fongbe verbs *fó* and *vɔ̀* 'finish' in the SVC are more grammaticalized than in the Saramaccan counterpart SVC, which does not allow them, since 'finish' implies a natural endpoint. Also, unlike Saramaccan, the Fongbe V2 is more grammaticalized in that it cannot be predicate-clefted (Lefebvre & Brousseau 2002). The 'finish' SVC in Saramaccan, like the comparison SVC, would appear in a normal diachronic context, hypothetically from Fongbe to Saramaccan, to be the reverse direction of grammaticalization; again, as with other SVCs, this is an example of greater semantic transparency in creoles. This difference in verbal properties may have been significant in the transfer of the *kabá* SVC; in a non-serial *kabá*-final clause the etymon *kaba* originally meant 'already', since grammaticalized as *kàà*. There is a Fongbe etymon for 'already', the anterior marker *kò* (p. 89), which occurs before the VP like other TMA markers. But even though its position is somewhat parallel to the English 'already', *kò* was not relexified as 'already', perhaps because in Fongbe it is also a TMA marker. It was the Fongbe verbs *fó* and *vɔ̀* 'finish' (p. 433) that were relexified in V2 position to mean both 'finish' and 'already', probably because these verbs lacked verbal properties relative to other Fongbe V2s and so could be interpreted as either verbs or adverbs. Incidentally, the 'finish' SVC in Saramaccan can also have a V2 *tapá* 'stop', 'cover', and the Fongbe 'finish' SVC can also have V2 *té* 'stop' (434).

In the early Saramaccan Wietz text the *kaba* SVC is attested, unlike the *pasa* SVC, and in Sranan it is attested in all the early Sranan texts as well (Arends & Perl 1995: 56). But although there are many clauses in the Wietz text ending with *kaba*, not all of these would be considered SVCs by the constraints on the modern *kabá* SVC. Of the 22 such clauses in Wietz, only 4 would be SVCs in modern Saramaccan, for example:

(5) teh dem bi begi di begi kaba
until they ANT pray the prayer finish
'(and) when they had prayed' (Wietz; Acts 4: 31)

The Bible translation here and throughout this article is from the Revised Standard Edition. The other 18 *kabá*-final clauses depict STATES, ACTIVITIES or ACHIEVEMENTS, for example; *hem dedde kaba* (he-*emph* dead finish) 'he … died' (Acts 2: 29).

In the Wietz text clauses ending with *kaba* can usually mean 'already', modern Saramaccan *kàà* 'already'. Interestingly, none of the 22 passages with *kaba* in Wietz are translated with the *kabá* SVC in the modern Saramaccan Bible (SIL 1991), even the 4 clauses in Wietz that could have the SVC; the SIL version of Acts 4:31 is *di de begi sö te de kaba* (when they pray so until they finish). For the Wietz *kaba* clauses in SIL there are 7 occurrences of *kàà* 'already' and 3 variations of the common phrase *té kabá* 'until finished'. In fact, in the entire SIL version of the Wietz text, Acts 1–13:33, there are no *kabá* SVCs at all. This SIL text has 36 occurrences of *kàà*, and 31 of variations of *té kabá*. In addition, in the 74-page narrative of Laurens Aboikoni there is only one *kabá* SVC (Aboikoni & Glock 1997:44), although there are numerous occurrences of *kàà* and *té kabá*. This confirms speaker intuition that the *kabá* SVC is dispreferred and rare. For comparison, Sebba (1987:73) considers the corresponding construction in Sranan with *kpa* 'finish' is not an SVC, with *kpa* instead being an adverb meaning 'already' rather than a V2.

In the Wietz text there are 2 non-serial alternative constructions, both of which depict ACTIVITIES, for example:

(6) di Sombre no kaba va takki kosi Oto va di santa Gamja
the person not finish for talk curse word of the holy place
'this man never ceases to speak words against this holy place' (Wietz; Acts 6:13)

In SIL there is one; *an kaba u fan* (not finish for speak) 'still saying (this)' (Acts 10:44). In the Aboikoni text, which has one *kabá* SVC, there are 7 non-serial alternative constructions with *kabá* (*u*). The less constrained *kabá u* construction is the more preferred, in texts and elicitation, and speaker intuition. But the *kabá u* construction is not a substrate feature. In Fongbe the aspectual verbs *fó* and *vɔ̀* do not occur before main verbs (Lefebvre & Brousseau 2002).

A likely transfer scenario would posit a pidgin 'finish' construction as two independent clauses, one with the matrix verb and the other with 'finish'. The transfer of the Fongbe 'finish' SVC would occur later and be ambiguous with both a serial 'finish' and a non-serial 'already' meaning.

2.1.3 *Directional SVCs*

The directional SVCs use V2 *gó* 'go' or *kó* 'come' to indicate the direction of the action of V1 as well as the action of 'going' or 'coming'. Non-serial alternatives to intransitive directional SVCs can be separate clauses, although the SVC use of V2 *gó* or *kó* is much more common and more acceptable to speakers. Non-serial alternatives to transitive directional SVCs can simply not use V2 *gó* or *kó* if the direction can be determined by context, although here too the SVC use is more common. But from the limited data in the early texts it would appear that in the late 18th century non-serial alternatives were much more common. The 1805 Wietz text has been compared to the modern SIL Saramaccan Bible (1991) in Kramer (2002), since both the Wietz (Arends 2002:127) and the SIL Bible (Glock p.c.) were written by missionaries in collaboration with speakers

of Saramaccan. Comparing the non-serial alternatives in every non-serial directional construction (i.e. without *gó* or *kó* as V2) in Wietz to the same passages in SIL, nearly every corresponding passage in SIL has a directional serial. For example, in Wietz there are 7 SVCs with *tjarri go/kom* (carry go/come) 'bring'/'take', while there are 12 occurrences of *tjarri* in single verb clauses. In the corresponding section of the SIL Bible, there are 73 SVCs with *tja go/ko*, and 16 occurrences of *tja* as a single verb. Similarly, in Wietz there are 4 SVCs with *manda go/kom* (send go/come) 'send (there)'/'send (here)', and 19 occurrences of *manda* as a single verb. In SIL there are 39 SVCs with *manda go/ko*, and 35 occurrences of *manda* as a single verb. From this data it seems that an earlier preference for the non-serial alternatives to directional SVCs has changed in the modern language to a preference for the SVCs. This change in the variation between the directional SVC, as well as other SVCs, and the non-serial alternatives since the 18th century suggests that the late increase in the frequency of use of the SVC option could be the result of late transfer.

Examples (7b.i) and (7b.ii) are from the Wietz text. Speakers say that *tjá* 'carry' can mean 'bring'/'take' with or without *gó* or *kó*, since 'carry' implies motion with the direction generally understood, but that the SVC with a directional verb is preferred:

(7) a. i. tótómbótí tjá wáta kó é
woodpecker carry water come *exp*
'Woodpecker brought water!' (Totomboti; folk tale)

ii. wā́ hógi tjá wā́ búnu
one terrible carry one good
'a bad thing brings a good thing' (proverb)

b. i. dem tjarri hem go na dorro
3pl carry 3sg go at door
'they brought him out' (Wietz; Acts 5:6)

ii. a tjarri dem na dorro
3sg carry 3pl at door
'he brought them out' (Wietz; Acts 5:19)

c. i. de tja dee lö soni dë ko
3pl carry DETpl type thing there come
'they brought those kinds of things' (Aboikoni & Glock 1997: 10)

ii. dee sëmbë tja gudu
DETpl person carry goods
'the people brought goods' (Aboikoni & Glock 1997: 9)

d. i. de manda wan biifi go na asindo opo
3pl send one letter go at Asindo Opo
'they sent a letter to Asindo Opo' (Aboikoni & Glock 1997: 13)

ii. kantoo manda di buka a hii dee köndë
office send DETsg mouth at all DETpl village
'they sent a message to all the villages' (Aboikoni & Glock 1997: 37)

e. i. a tjá dí mɔ́ni gó a wowójo
3sg carry the money go to market
'he took the money to market'

ii. a tjá dí mɔ́ni hẽ́ a gó a wowójo
3sg carry the money and 3sg go to market
'he took the money and went to market'

The Fongbe directional SVC, like all simple (Saramaccan type) SVCs, can be paraphrased with a coordinate construction:

(8) kɔ̀kú sɔ́ kòklô ɔ́ bó yì àxì mɛ̀ Fongbe
Koku take chicken DET and go market in)
'Koku took the chicken and went to the market' (Lefebvre & Brousseau p. 400)

This option, of course, is available for the directional SVCs in Saramaccan, in (7e.ii). As another alternative in Fongbe, there are transitive verbs depicting movement with a locational object:

(9) é tɛ́ è àyì Fongbe
3sg knock down 3sg ground
's/he knocked her/him flat (to the ground)' (Lefebvre & Brousseau p. 239)

Unlike Saramaccan, Fongbe simple SVCs do not allow predicate clefting of V2 (da Cruz 1994; Ndayiragije 1992), although Lefebvre and Brousseau (2002) report some speakers do allow it (p. 406). The V2s of Fongbe simple SVCs, then, do not have the verbal clefting characteristic of V1s and non-serial verbs, and could therefore be seen as more grammaticalized than their Saramaccan counterparts. And as with other SVCs, there appears to be an increase in the frequency of use of the SVC compared to the non-serial alternative since the early texts, suggesting late transfer.

2.1.4 *Complementizer constructions*

Complementation with *táa*, grammaticalized from *táki* 'talk', would appear to be a SVC (Veenstra 1996: 93). The etymon *táa* appears to occur as a verb meaning 'talk', although unlike *táki* it may not be followed by a complementizer; **a táa táa* (3sg talk COMP), **a táa dá mí táa* (3sg talk give 1sg COMP). But *táa* can appear to be a verb:

(10) de taa wë sö de ko aki
3pl 'say' EMPH so 3pl come here
'they said they had come' (Aboikoni & Glock 1997: 37)

Speakers say, however, that these occurrences of *táa* are really *táki táa* with the understood elimination of *táki* 'talk', so that *táa* would not necessarily be a verb. In fact *táa* is a complementizer: it lacks V2 verbal characteristics, such as aspect marking and predicate clefting, so that constructions such as (11a) are not SVCs (Veenstra 1996: 157). As

in Sranan (Plag 1993: 40), indirect speech may occur in Saramaccan with or without the complementizer, as in (11b):

(11) a. a táki táa a ó kó amãjá
3sg talk 'say' 3sg FUT come tomorrow
'he said that he'll come tomorrow'
b. a táki a ó kó amãjá
3sg talk 3sg FUT come tomorrow
'he said he'll come tomorrow'

There are two analyses for the Fongbe complement introducing construction, perhaps representing two dialects; one is that it is a SVC with V2 ɖɔ̀ 'say', and the other is that ɖɔ̀ is a complementizer (Lefebvre & Brousseau 2002: 115). For many speakers ɖɔ̀ does not occur after the main verb ɖɔ̀ 'say' due to a repeated morph constraint; *kɔ̀kú* ɖɔ̀ *(ɖɔ̀)bàyí wá* 'Koku said (that) Bayi came' (p. 116).

In the early Saramaccan texts there is no occurrence of a *takki* complementizer equivalent to the modern *táki táa*. Complement clauses use the preposition *va*, as *a takki va dem … go* 'he said that (they) should go' (Wietz Acts 5:34), or simply a null complementizer. This is also the case in the earliest texts in Sranan (Plag 1993: 62). In Sranan the *taki* complementizer then basically developed from an apparent SVC V2 for speech act, cognitive and perception verbs, into a grammaticalized complementizer (p. 79). In Sranan and perhaps Saramaccan it would appear that the complementizer strategies developed and then grammaticalized late. The fact, however, that the eventual complementizer construction matches the substrate could not be coincidence. The constant importation and increasing percentage of new speakers of Gbe languages (Migge 1998: 130) meant a constant substrate influence during the generation before the time of the complementizer development. The complementizer construction, then, is a good candidate for late transfer, perhaps from both the Fongbe SVC and the non-serial complementizer construction.

2.1.5 *Instrumental SVCs*

A stylistic use of the instrumental SVC, and common for other SVCs as well, is its use in information packaging as a kind of scene-setting for subsequent event depictions, a marked function. In the single occurrence of an instrumental SVC in the 1805 Wietz text, for example, the entire event of the SVC could be seen as scene-setting:

(12) teki di bigi Areen-Jakkiti tappa ju Skin en kom flakka mi
take the big (rain?)-jacket cover you skin (body) and come follow me
'Wrap your mantle around you and follow me' (Wietz; Acts 12:8)

This is then followed by a series of single verb clauses depicting events. Similarly, in modern Saramaccan a depiction of clearing a path would use a non-serial construction like (13a), but for setting the scene for subsequent events the SVC would be used, in (13b):

(13) a. a kóti dí pási ku dí ufángi
3sg cut the path with the machete
'he cut the path with the machete'
b. a téi dí ufángi kóti dí pási
3sg take the machete cut the path
'he took the machete and cut the path'

This is one way, then, that the instrumental and other SVCs are used relative to the non-serial alternatives. Another way Saramaccan instrumental SVCs differ from non-serial alternatives, and in fact are marked relative to them is that, like Saramaccan shared object SVCs (Kramer 2002: 287), presented below in 2.1.6, they rarely occur and are generally unacceptable to speakers in ACTIVITY depictions. To be acceptable, the affected arguments, for example, must be specific; **a téi dí ufángi kóti páu* ('he took the machete and cut trees').

Comparing the Wietz text with the same passages in the modern SIL Bible (Acts 1–13:33), there is the one instrumental SVC in Wietz, above, and 10 instrumental SVCs in SIL (including the one in Acts 12:8). This extremely small sample is significant only in that it parallels the ratios in comparisons of the two texts with the other SVCs, indicating an increase in the use of SVCs since the early texts.

In (14), below, examples (14a.iii) as well as (14a.ii) are non-serial alternatives to the SVC in (14a.i). Example (14c) shows constraints on constituents between the SVC verbs, discussed below:

(14) a. i. a téi dí góni súti dí píngo
3sg take the gun shoot the pig
'he took the gun and shot the pig'
ii. a súti dí píngo ku dí góni
3sg shoot the pig with the gun
'he shot the pig with the gun'
iii. a téi dí góni hɛ́ a súti dí píngo
3sg take the gun then 3sg shoot the pig
'he took the gun and shot the pig'
b. i. a fufúu dí góni súti dí píngo
3sg steal the gun shoot the pig
'he stole the gun and shot the pig'
ii. a fufúu dí góni hɛ́ a súti dí píngo ku ɛ́
3sg steal the gun then 3sg shoot the pig with 3sg
'he stole the gun and shot the pig with it'
c. a téi / fufúu / kísi / (*sáka) dí góni wáka gó súti dí píngo
3sg take / steal / catch / put down the gun walk go shoot the pig
'he took / stole / caught / (*put down) the gun, walked, went and shot the pig'

In Saramaccan instrumental SVCs the V1 is commonly *téi* 'take', but like the V2 of this SVC, it may be any verb from the open class of dynamic object-affecting transitive verbs, except that it is pragmatically constrained to depict, sometimes figuratively, 'holding' the object as an instrument, as in (14c). The occurrence of unergative verbs between the two transitive verbs, like *wáka* in (14c), is determined by the semantics of V1 regarding 'holding'; actions like 'taking', 'stealing' or 'catching' can continue while 'walking'. An action like *sáka*, however, which denotes 'lowering and letting go' could not continue while 'walking', and so could not allow any instrumental use after 'walking'.

In contrast to Saramaccan, the V1 of the Fongbe instrumental SVC is from a closed class; *sɔ́ /zé* 'take' (Lefebvre & Brousseau 2002: 415), and so unlike other SVCs, the V1 is the *serial item*, the more grammaticalized verb in an SVC. In Saramaccan, then, the V1 is more verbal and less a grammaticalized serial item than its Fongbe counterpart. In fact, the V1 in Saramaccan may not be a serial item, as it is V2 that is constrained in tense marking. Fongbe has non-serial instrumental constructions, like (14a.iii), using the *bó* 'and' coordinate construction (p. 416), or an instrumental PP construction, like (14a.ii), with *kpó* 'with' (302). Because there are similar constructions in English, there would be no motivation for incorporating an instrumental SVC into an MIC, as it would be nonessential for the limited functions of an English-Fongbe pidgin. It would therefore have entered the creole after the pidgin stage, and so would have transferred without being affected by the pidgin filter. And, as with other SVCs, there appears to be an increase in the frequency of use since the early texts, suggesting late transfer.

2.1.6 *Shared object SVCs*

Shared object serials in Saramaccan are used to introduce or reactivate arguments with full NPs (Kramer 2002: 603), as they rarely occur with pronominal shared objects. In naming arguments, shared object SVCs, like instrumental SVCs, often have a scene-setting function for subsequent events. One discourse function of shared object SVCs, then, is to introduce or to bracket episodes (pp. 290–295). For example, in a story about tapir catching fish by eating a poison *ndeku* root, a shared object SVC sets the scene for the subsequent series of events depicted by single verb clauses (p. 290):

(15) a tá njã́ dí ndekú butá nɛ̃ɛ̃́ bɛ́ɛ
3sg CONT eat the *ndeku* put in-his belly
'he ate the *ndeku*, put it in his belly'

The V2 phrase *butá nɛ̃ɛ̃́ bɛ́ɛ* 'put in his belly' adds almost no information to the V1 phrase, but the resulting shared object SVC provides the background, as tapir defecates in the creek and fish are poisoned. For an example of bracketing, in a story about making cassava bread for gravediggers, a shared object SVC occurs at the beginning and another occurs at the end of a sequence of events depicted by single verb clauses (Glock 1972: 60–61, cited in Kramer 2002: 293–294):

(16) de bì gó díki kasába tjá kó
3pl ANT go dig cassava carry come
'they went and dug cassava, and brought it'

The preparation includes grating, pressing, cooking, drying, pounding and baking. The shared object SVC bracketing the end of the sequence is then:

(17) de jasá kasába fu baákama féni tjá gó
3pl bake cassava for gravedigger find carry go
'they baked cassava bread for the gravediggers to find and take (with them)'

In the Wietz text the introducing function is seen with two nearly identical shared object SVCs:

(18) a. a teki di Moni putta na Futtu va dem Apostel
3sg take the money put at foot of the(pl) apostle
'he brought the money and laid it at the apostles' feet' (Wietz; Acts 4:36)
b. di otro pisi a tjari putta na futu va dem Apostel
the other piece 3sg carry put at foot of the(pl) apostle
'(he) brought (the other part) and laid it at the apostles' feet' (Wietz; Acts 5:2)

The bracketing function is seen in Acts 5:2–9, for two dialogues by Peter on a single topic, referencing the two other participants, where 5:2 and 5:9 have bracketing shared object SVCs:

(19) dissi tjari ju Manu go beri
this carry your husband go bury
'(those) who have buried your husband' (Wietz; Acts 5:9)

In the same passages in the SIL Bible the dialogues are similarly introduced and bracketed, but with five other shared object SVCs bracketing further divisions. The comparison of these passages is parallel to a comparison of all the passages; in Wietz there are four shared object SVCs, while in SIL there are 114, an increase of 96.5%. As with other SVCs, the late increase in usage of the SVC variant suggests a late transfer, at a time, of course, previous to the early texts.

In the comparison of the two texts there is another difference. In Wietz there are paratactic constructions that could be either conjoined clauses or shared object SVCs in modern Saramaccan. For conjoined clauses, gapping, allowed in English, is not accepted in modern Saramaccan; both conjunction and subject pronoun are required. On the other hand, pleonastic pronouns are not acceptable in SVCs in modern Saramaccan. In Wietz there is an example:

(20) (dem) kom findi hem, tjarri hem na dorro, dem beri hem
(3pl) come find 3sg, carry 3sg to door, 3pl bury 3sg
'the young men … found her … and they carried her out and buried her'
(Wietz; Acts 5:10)

In the same passage in SIL there is a directional SVC as part of a shared object SVC:

(21) de … hopo ën tei tja go bei
3pl raise 3sg take carry go bury
'they carried her out and buried her' (SIL; Acts 5:10)

There are 16 such paratactic constructions in the Wietz text. The change away from paratactic constructions and toward greater use of shared object SVCs suggests that this stylistic variation was resolved in favor of substratal influence, perhaps, as argued by Sankoff (1994), beginning by later generations' retention of substrate features for use as the creole language expands to take on more functions, and continuing by language identity resulting from Saramaccan independence gained in 1762 (Le Page & Tabouret-Keller 1985, cited in Kramer 2002: 605), discussed in Section 3 below. Without the paratactic constructions in modern Saramaccan, then, the non-serial alternatives to shared object SVCs are only coordinate clauses or single verb clauses. In (22c) the unacceptability of *sáka* is from Veenstra (1996: 142):

(22) a. i. a téi dí búku butá a dí táfa
3sg take the book put at the table
'he took the book and put it on the table'
ii. a téi dí búku hɛ̃ a butɛ̃ɛ̃ a dí táfa
3sg take the book then 3sg put-3sg on the table
'he took the book and put it on the table'
b. i. de hái dí bóto butá a téla
3pl haul the boat put on shore
'they hauled the boat ashore'
ii. de hái dí bóto hɛ̃ de butɛ̃ɛ̃ a téla
3pl haul the boat then 3pl put-3sg on shore
'they hauled the boat and put it ashore'
c. a téi / fufúu / kísi / (*sáka) dí búku wáka go butá a dí táfa
3sg take / steal / catch / put down the book walk go put on the table
'he took/stole/caught/(*put down) the book, walked, went and put it on the table'
d. i. a téi dí báta boóko
3sg take the bottle break
'he took the bottle and broke it'
ii. a téi dí báta hɛ̃ a boóko ɛ̃
3sg take the bottle then 3sg break 3sg
'he took the bottle and broke it'
e. i. a náki dí báta boóko
3sg hit the bottle break
'he hit the bottle and broke it'

ii. a náki dí báta hɛ̃ a boóko ɛ̃
3sg hit the bottle then 3sg break 3sg
'he hit the bottle and broke it'

f. i. mi ló u fáa déé páu túwɛ
1sg love for fell the(pl) tree throw
'I love to fell the trees'

ii. mi ló u fáa páu
1sg love for fell tree
'I love to fell trees'

Both V1 and V2 of the Saramaccan shared object SVC are from an open class; dynamic object-affecting transitive verbs (Kramer 2002: 188), where effect includes change of location (Kramer p. 227; Hale & Keyser 1987). In effect the shared object SVC is ergative marking, an unexpected feature in creole languages (McWhorter 2002: 9); the NP1 is always the Agent, and the Patient NP2 shared object is also the subject of intransitive serial verbs subsequent to the shared object SVC (Kramer 2002: 218). Concerning unergative verbs within the SVC, the constraint on instrumental SVCs, in (14c), also applies to shared object SVCs, in (22c); the action of V1, 'taking' or 'stealing' but not 'putting down', must be able to continue while 'walking'. The difference between (14c) and (22c), of course, is that in the shared object SVC the affected object of V2 is NP2, while in the instrumental SVC it is NP3. An additional difference in Fongbe is that for instrumental SVCs the closed class verb, or serial item, is V1, and in the resultative shared object SVC it is V2.

The open class selection of both V1 and V2 verbs in this SVC in Saramaccan contrasts with Fongbe, where the number of V2 verbs appears to be limited to about a dozen (Lefebvre & Brousseau 2002: 431–432). Saramaccan shared object SVCs rarely occur in ACTIVITY depictions in texts or spontaneous narrative, and they are dispreferred in elicitation (Kramer 287). In the Wietz text there were no ACTIVITY shared object SVCs, and in the SIL text only 2.6% of shared object SVCs were ACTIVITY depictions. So like other SVCs in Saramaccan, the shared object SVC is constrained relative to the non-serial alternatives. From the examples in the literature it appears that the Fongbe shared object SVCs tend not to occur in ACTIVITY depictions as well. In Saramaccan there is some degree of idiomaticity, especially with change of location verbs like *túwɛ* 'throw (to)' and *púu* 'push (from)', such as *fáa dí páu túwɛ* 'fell the tree', in (22f.i), where V1 *fáa* 'to fell' more than *kóti* 'cut' or *sã́* 'saw' selects V2 *túwɛ* which depicts the action as resultative, but the absence of V2 does not change the basic meaning. The non-serial clause does, however, allow an ACTIVITY depiction, in (22f.ii). There are also institutionalizations (McWhorter 2004: 88), whose compositionality is accessible, like *táki túwɛ* (talk throw) 'to talk excitedly'. The meanings of the verbs in Saramaccan shared object SVCs are transparently compositional. This contrasts with Fongbe, where a large proportion of the shared object SVCs are lexicalizations with specialized meanings (Lefebvre & Brousseau 2002: 431). The verbs in Saramaccan shared object SVCs retain their individual meanings more, and are there-

fore generally less grammaticalized than the verbs in Fongbe shared object SVCs. The more prototypically verbal items in Saramaccan shared object SVCs allow for the open class selection of V1 and V2, which is also a characteristic of SVC-like paratactic constructions in Ndyuka (Migge 1998: 154). The shared object SVC is like the instrumental SVC, then, in that except for the tense marking constraint on V2 it is basically an SVC without serial items; it is an SVC in that two or more verbs share an NP argument and depict a single event, but without one of the verbs being grammaticalized and selected from a small closed set, with the loss of some prototypical verbal characteristics such as those described by Migge (1998: 223–240). The greater verbal prototypicality means that the verbs in a shared object SVC are more like main verbs, which allows a ready comparison to non-serial alternatives. The Saramaccan shared object SVC is less grammaticalized as well in that it allows clefting of both verbs, where the Fongbe shared object SVC allows only clefting of V1 (Lefebvre & Brousseau 2002: 432). And, more than with other SVCs, there is an increase in the frequency of use, namely 96.5%, in the Wietz/SIL comparison, suggesting late transfer.

2.2 SVCs in the Early Saramaccan Texts

2.2.1 *GOAL/BEN SVCs and ditransitives*

There are greater differences between the early texts and the modern language for GOAL/BEN SVCs than for other SVCs. In modern Saramaccan GOAL/BEN depictions there is no PP non-serial alternative, as there is in the English dative shift. There is a PP alternative in the early texts, however. In modern Saramaccan there is a ditransitive non-serial alternative to the GOAL/BEN SVC, in (23a.ii), (23b.ii) and (23c.ii), but there are very few ditransitive verbs, possibly 6 or 7 depending on the speaker. For GOAL/BEN depictions with other verbs in modern Saramaccan, then, the GOAL/BEN SVC is the only option, and so is not a stylistic alternative.

The GOAL/BEN SVC with NP2 is a type of shared object SVC, where change of location is considered an effect (Hale & Keyser 1987), as well as change of possession. The V2 is one of the ditransitives, which as with other shared object SVCs depicts the endpoint of the change. An affected argument, then, does not follow V2, in (23d.ii), (23e.ii) and (23f.ii):

(23) a. i. a dá dí mɔ́ni dá mí
3sg give the money give 1sg
'he gave the money to me'

ii. a dá mí dí mɔ́ni
3sg give 1sg the money
'he gave me the money'

b. i. a paká dí mɔ́ni dá mí
3sg pay the money give 1sg
'he paid the money to me'

ii. a paká mí dí mɔ́ni
3sg pay 1sg the money
'he paid me the money'

c. i. a léi dí fóótóo dá mí
3sg show the photo give 1sg
'he showed the photo to me'

ii. a léi mí dí fóótóo
3sg show 1sg the photo
'he showed me the photo'

d. i. a téi dí fóótóo léi mí
3sg take the photo show 1sg
'he took the photo and showed it to me' (Veenstra 1996: 86)

ii. *a téi mí léi dí fóótóo
3sg take 1sg show the photo
(he took me and showed me the photo)

e. i. a téi dí mɔ́ni paká mí
3sg take the money pay 1sg
'he took the money and paid me'

ii. *a téi mí paká dí mɔ́ni
3sg take 1sg pay the money
(he took me and paid me the money)

f. i. a fení dí mɔ́ni paká mí
3sg find the money pay 1sg
'he found the money and paid me'

ii. *a fení mí paká dí mɔ́ni
3sg find 1sg pay the money
(he found me and paid me the money)

Saramaccan is like Fongbe in having only a few ditransitives, Saramaccan has 6 or 7, and Fongbe probably has 3 or 4 (Lefebvre & Brousseau, 254, 445). For other verbs, then, an SVC with a ditransitive V2 is used. In Saramaccan, above, and in Fongbe the GOAL/BEN argument cannot be the shared argument (Lefebvre & Brousseau 2002: 411), since to be 'shared' an argument needs to be affected by each verb. Unlike modern Saramaccan, in Fongbe there are non-serial alternatives for GOAL/BEN depictions using prepositions (pp. 314, 465). But there are prepositional alternatives in the early Saramaccan texts. The stylistic distribution of the serial and non-serial constructions suggests a typically creole sensitivity to the semantics of *da* 'give' and the preposition *na* 'to'. Prepositional alternatives to GOAL/BEN SVCs exist in early Sranan as well (Arends & Perl 1995), with what would appear to be a similar distribution. Because of their similarity to English as well as Fongbe, the prepositional alternatives to the GOAL/BEN SVC probably transferred in the MIC stage. Looked at from another point of view, McWhorter (1996: 39–44) considers that a lack of SVCs

with prepositional functions in an early pidgin stage may have enhanced the transfer of prepositions.

2.2.2 *GOAL/BEN SVCs in early texts*

In modern Saramaccan GOAL/BEN is expressed by SVCs with V2 *dá* 'give'. In the earliest texts, however, GOAL/BEN was expressed by an SVC with V2 *da* or *gi* 'give', or by a non-serial construction with the prepositions *na* 'at'/'to', or *va* 'of'/'for', for BEN only. There is a substrate source for this *give/to* alternation, with the Fongbe verb *ná* 'give' and the preposition *nú* 'at', where there is a PP non-serial alternative in Fongbe (Lefebvre & Brousseau 2002: 473):

(24) kɔ̀kú nyà àkwɛ́ nú àsíbá Fongbe
Koku loan money to Asiba
'Koku loaned money to Asiba'

There are West African languages where the verb 'give' and the preposition 'for' are homophonous (p. 321), and, regarding lexical conflation in creolization (Kihm 1989), there is segmental identity between Fongbe *ná* 'give' and early Saramaccan *na* 'at'/'to' (Enoch Aboh p.c.), but the fact that the SVC exists in early Saramaccan and Sranan indicates the verb was not confused with the preposition.

In early Saramaccan the SVC with *da* occurs as a stylistic variant where the notion of 'giving' would be appropriate. This sensitivity to the semantics of a V2 is not unexpected in a creole, where semantic transparency is greater than in a source language (McWhorter 2002). As 'giving' entails the reciprocal notion of 'receiving', the SVC with *da* occurs where there would be the expectation of a volitional response. Thus *takki da* would be 'talk with' while *takki na* would be 'talk to'. The corresponding SVC in early Sranan, with V2 *gi* 'give' has a similar distribution, for example; *takki gi* where a volitional response is expected (Arends & Perl 1995: 279) and *takki na* for a command (p. 172). Because of the expected response, then, the SVC in early Saramaccan tends to occur mainly with GOAL/BEN participants who are of equal or lower social status than the Agent participant. As with other SVCs and alternatives, this constraint does not apply to the non-serial alternative constructions with *na* or *va*, or ditransitives. Thus in (25), a lie may be 'given' to a person, who would be expected to acknowledge and/or respond, but only 'presented' to God:

(25) ju no bi lögede <u>da</u> Sombre, ma ju lögede Gado.
2sg NEG ANT lie give person but 2sg lie God
'you have not lied to men but to God.' (Wietz; Acts 5:4)

There are apparent exceptions in passages that mock the power differential:

(26) dem meki sondi <u>da</u> di Gado va dem
3pl make thing give the God of 3pl
'they offered a sacrifice to the idol' (Wietz; Acts 7:41)

After seeing the offering, God mocks the Jews with the similarity to their practices. The mocking passages are Acts 7:42, 47, 49, and the mocking of Saul in 9:21:

(27) hu speri Hosso Unu sa meki da mi
what type house 2pl FUT make give 1sg
'What house will you build for me . . .' (Wietz; Acts 7:49)

The distribution of V2 *da* in Wietz holds for the maroon letters (Arends & Perl 1995: 383–387), which have GOAL/BEN SVCs with *da*, and also with *gi* 'give', in (28a–d). As in the Wietz text, the non-serial PP construction with *na* 'at'/'to' is not constrained by the semantics of 'giving', and occurs where *da* does not, in (28e–f):

(28) a. a takki da wi . . . ote va Jesus Christus da wi
3sg talk give 1pl word of Jesus Christ give 1pl
'He (missionary) told us the story of Jesus Christ . . .' (letter no. 2 part 1)
b. wi takki grang oudi gi oene
1pl talk big greeting give 2pl
'we (our brothers) say goodbye to you (missionary)' (letter no. 2 part 1)
c. Mi ben takki gi dem
1sg ANT talk give 3pl
'I told them (Saramaccans)' (letter no. 5)
d. so mi no minga va takki da dem
so 1sg NEG deny of talk give 3pl
'so I couldn't help but tell them (Saramaccans)' (letter no. 5)
e. wi tan kreh na Helpiman
1pl stand cry to help-man
'we are crying out to our Savior' (letter no. 2 part 1)
f. mi beggi na Helpiman
1sg beg to help-man
'I begged the Savior' (letter no. 2 part 1)

While *da* entails sharing and reciprocity, *na* tends towards more to depict remoteness. *Da* and *na* often function together highlighting this difference, such as calling attention with *na*, and then giving instructions with *da*, in (29a–b) and (29c). The use of *na* in (29a) instead of *da* is not unexpected for calling out a name. The use of *da* in (29b) is also not unexpected, as giving instructions generally involves a volitional response. But the *na* in (29c) is used in giving instructions in a different context; unlike (29a) where the addressee is a disciple, the addressee in (29c) is Saul, who is accused of persecuting Jesus, the addresser, so the statement is more an order than an instruction. When Saul gets to the city he will no longer be a persecutor, so he will be instructed by peers, with *da*:

(29) a. Massra takki na hem, na wan feesi: Ananias!
master talk to 3sg to one face Ananias
'The Lord said to him in a vision, Ananias!' (Wietz; Acts 9:10)

b. Massra takki da hem; hoppo nappeh go na di Paasi …
master talk give 3sg rise to-place go to the street
'… the Lord said to him, 'Rise and go to the street …' (Wietz; Acts 9:11)

c. Massra takki na hem: hoppo tan nappeh, go na dindru di Foto;
master talk to 3sg rise stand to-place go to inside the fort
'and he said … rise and enter the city, …'
alla dem sa takki da ju, ondi ju sa du
all 3pl FUT talk give 2sg what 2sg FUT do
'… and you will be told what to do' (Wietz; Acts 9:6)

Another example of the use of *da* compared to *na* is from Wietz (Acts 7:2 and 7:30). In Acts 7:2:

(30) Gado … bi sori hem srefi da wi Tatta Abraham
God … ANT show 3sg self give our father Abraham
'God … appeared to our father Abraham' (Wietz; Acts 7:2)

Then God talks to Abraham; *A takki da hem* '(He) said to him' (Acts 7:3). The use of *da* indicates that Abraham will be given a good thing, namely directions to the Promised Land. In Acts 7:30, on the other hand, Moses is being admonished:

(31) Engel … sori hem srefi na Moses
angel … show 3sg self to Moses
'an angel appeared to (Moses)' (Wietz; Acts 7:30)

Then God talks to Moses:

(32) tongo va Massra a takki na hem
tongue of master 3sg talk to 3sg
'The voice of the Lord came (it said to him)' (Wietz; Acts 7:31)

The use of *na* indicates a less reciprocal interaction than with Abraham. Moses is afraid to look, and trembles, as God tells him to take off his shoes because he is on holy ground.

Because *takki na* denotes 'talking to', unlike *takki da* 'talking with', there wouldn't be the volitional acceptance with *na* that is associated with *da*. An expected use of *na*, then, is:

(33) dem takki na hem; ju Lau
3pl talk to 3sg 2sg crazy
'They said to her 'You are mad'.' (Wietz; Acts 12:15)

The V2 *da* has a more personal connotation than *na*, in (34a–c):

(34) a. meki wi takki da dem bunbun,
make 1pl talk give 3pl good-good
'let us warn them …'

va dem no takki na wan Sombre morro va di nem.
for 3pl NEG talk to one person more of the name
'… to speak no more to any one in this name.' (Wietz; Acts 4:17)

b. kaba a preki da hem va Jesus
then 3sg preach give 3sg of Jesus
'and … he told him the good news of Jesus.' (Wietz; Acts 8:35)

c. a preki di Evangelium na tulu Foto
3sg preach the gospel at all fort
'he preached the gospel to all the towns …' (Wietz; Acts 8:40)

While *na* depictions are generally locative, V2 *da/gi* constructions are generally GOAL/BEN depictions:

(35) a. a putta hem da dem, na dem Feesi libi libi
3sg put 3sg give 3pl at their face live live
'he presented her (to the saints and widows) alive' (Wietz; Acts 9:41)

b. kaba a teki di Moni putta na Futtu va dem Apostel.
finish 3sg take the money put at foot of the(pl) apostle
'and (he) brought the money and laid it at the apostles' feet.' (Wietz; Acts 4:36)

c. na hem mi tan sindi di briffi gi hem
to 3sg 1sg stand send the letter give 3sg
'to him I send this letter' (letter no. 2 part 1)

But some occurrences of *da* or *na* appear to be blocked for semantic reasons. For example, *konda* 'relate' only occurs with *da*, in 5 occurrences in Wietz, while *leri* 'teach' only occurs with *na*, in 2 occurrences. There are too few occurrences to be significant, but the semantics of 'relate' would imply 'giving', while 'teach' would imply 'imparting'. The V2 *da* does not occur with V1 *da* (*kom da da* 'related to' in 10:8 is probably a mistranscription of *konda da* 'relate to'), perhaps because the notion of 'giving', and therefore GOAL, is denoted by V1 *da* 'give', in (36):

(36) a. bika a no da di nem na Gado
because 3sg NEG give the name to God
'because he did not give God the glory.' (Wietz; Acts 12:23)

b. ju tan da sondi na Poman
2sg stand give thing to poor-man
'you give things to the poor' (Wietz; Acts 10:4)

Similarly the pattern with *da* is seen in BEN constructions, in (27) and (37a), where the V2 *da* would be appropriate for the BEN role. The use of *na* as BEN would be unexpected, in (37b–c), but *na* is not affected by the semantics of *da*. When used with *du* 'do', *na* depicts a BEN role, but there are too few occurrences to be significant; there are no occurrences of *du* with *da*, and only 3 occurrences of *du* with *na* in Wietz, and 1 in the maroon letters. Interestingly, there is a BEN interpretation with *na* in *na dem*

liba (at them above), meaning 'on them' (in Acts 8:16), but 'for them' (in Acts 8:15), below in (37d):

(37) a. Ma Engel … jabri dorro va Dongro Hosso da dem na neti
but angel open door of dark house give 3pl at night
'But at night an angel … opened the prison doors (for them)' (Wietz; Acts 5:19)

b. a bi du na di aki dissi Unu si … va a bi kom bunn
3sg ANT do at the here this 2pl see for 3sg ANT come good
'he has made this man strong whom you see …' (Wietz; Acts 3:16)

c. biegie sannie Gado ben du na mi …
big thing God ANT do at 1sg
'big things God has given me' (… done for me) (letter no. 2 part 2)

d. dem begi na dem liba
3pl pray at 3pl above
'they prayed for them' (Wietz; Acts 8:15)

The preposition *va* 'of'/'for' may also take a BEN object. As with *na*, there is often an impersonal connotation, but like *na* it's occurrence is not necessarily excluded by the distribution of *da*. Similar to *na … liba* 'for …', in (37d), there is also a common expression *va … heddi* (for … head) 'for … sake'. In modern Saramaccan there is the expression *gãã́ tangí fií* (*fu+í*), 'big thanks for you':

(38) a. unu membre unu srefi bunbun va dem Sombre
2pl remember 2pl self good-good for the(pl) person
'take care what you do with these men' (Wietz; Acts 5:35)

b. Unu begi Massra va mi heddi
2pl pray master for 1sg head
'pray for me to the Lord' (Wietz; Acts 8:24)

c. hu menni penna a musu jam va mi nem heddi
how many pain 3sg must eat for 1sg name head
'how much he must suffer for the sake of my name' (Wietz; Acts 9:16)

d. di dé Jamjam va mi Schee
the COP eat-eat for 1sg soul
'It is food for my soul' (letter no. 5)

It appears from the early texts that there was variation between the SVC with V2 *da* and non-serial alternatives with *na* or *va*, but these early texts do not provide a large enough body of data to make accurate predictions on the conditions for variation. The largest sample in Wietz is for the variation between the SVC *takki da*, with 29 occurrences, and *takki na*, with 33 occurrences. From this sample, though, it would seem that V2 *da* is consistently transparent in depicting the entailments of 'giving', and that this characteristically creole transparency can be seen in depictions with V2 *da* and other verbs.

The conditions for variation are often difficult to determine, even with variation in modern Saramaccan. What is important, however, is that there is variation, and that there is an increase in the frequency of occurrence for the SVC variant. For the variation between SVCs and non-serial alternatives in Saramaccan, it would appear that the use of SVCs has been increasing since the 18th century, and in the case of *takki da* has replaced the non-serial alternative altogether.

2.2.3 *GOAL/BEN SVCs in Maroon Spirit Language*

Bilby (1981, 1983, 1992) describes a language in Jamaica currently used for religious purposes but once a language for everyday use among descendants of Maroons. Referred to as Maroon Spirit Language, it shares numerous features with the Surinamese creoles not found in other Caribbean English creoles, including Jamaican. McWhorter (1996: 123–125) argues that the shared features indicate a common origin with Sranan, namely a fully developed Sranan creole brought to Jamaica by slaves transported from Suriname by the British, starting in 1671. Smith (2002: 133–134) argues that the origin of both Maroon Spirit Language and Sranan is an English pidgin common to the Caribbean, and that in the English Caribbean, except in the remote countryside of Jamaica and in Suriname, the pidgin was replaced by a more acrolectal creole. SVCs in Maroon Spirit Language would figure in this debate, if indeed SVCs were to be creole but not pidgin characteristics in the Atlantic English creoles. There are a few SVCs in Maroon Spirit Language in the limited data. A directional SVC matches Sranan and Saramaccan:

(39) cha in go na da bigi pre, kya in go na indi Maroon Spirit Language
carry 3sg go to the big place, carry 3sg go to inside
'take him to that grave, put him inside' (Bilby 1992: 9).

This matches the Saramaccan *tjá gó* 'carry go' SVC, and is interesting in that it shows the influence of basilectal Jamaican in *kya* 'carry', alongside Maroon Spirit Language *chai* 'carry' (Bilby 1983: 54). This directional SVC is not significant in the debate on Maroon Spirit Language influence on Jamaican because it is common to all the Afro-American creoles, including those lexified by English, French and Spanish/Portuguese (Alleyne 1980: 91). Similarly, Maroon Spirit Language has the GOAL/BEN SVC corresponding to the Saramaccan SVC *táki dá* (talk give) 'talk to' (Alleyne 2000: 128). This GOAL/BEN use of V2 'give' in SVCs is also common to all the Afro-American creoles (Alleyne 1980: 94). Bilby (1983: 48) notes that Maroon Spirit Language uses the preposition *na* 'at/to' in *tak na mi* 'talk to me', a construction and use of *na* that does not occur in Jamaican. Maroon Spirit Language also uses *tu* 'to' in *taak tu yu* 'talk to you' (p. 70), as in Jamaican *taak tu mi* 'talk to me' (p. 48). For a BEN role, *fi* 'for' is used; *peak fi mi* 'speak for me'. There is no mention of any use of a *taki ji/gi* 'talk give' SVC, which perhaps occurs somewhat rarely, and must be a stylistic variant with *taak tu* and *tak na*. In modern basilectal Jamaican the 'talk give' SVC occurs rarely as well, only where 'giving' is implied: *taak to* 'talk to', but *pap tuori gi* 'tell stories (lie) to' (Alleyne 2000: 127), a distribution similar to that of *takki da* and *takki da* in the early Saramac-

can texts, in 2.2.2 above. Of interest here is the variation between *takki na* and *takki gi* in the 1765 Sranan text, which also parallels the usage in the early Saramaccan texts: *anno takki na mi* 'he didn't tell me' (Arends & Perl 1995: 218) for information on the location of an evil manager, but *takki gi hem* 'tell him' (p. 229) when telling slaves to take the evil manager away, and *takki gi hem* 'tell him' (p. 232) when requesting rum for a party for slaves.

The GOAL/BEN SVC with V2 'give' is then, like other SVCs, common to the other Caribbean English creoles and therefore not significant in the debate on its influence on Jamaican. Maroon Spirit Language could be the oldest form of Jamaican (Alleyne 2000: 128; McWhorter 2000: 86), in which case the GOAL/BEN SVC in basilectal Jamaican would be relics of an earlier form of the language. Or, Maroon Spirit Language could be a remnant of an earlier pidgin, replaced by an acrolectal Jamaican, in which case the GOAL/BEN SVC in Maroon Spirit Language would be the result of contact with the acrolectal Jamaican, like other influences such as *kya* 'carry', above. The similarity of distribution for 'talk to' and 'talk give' GOAL/BEN SVCs between basilectal Jamaican and early Sranan and Saramaccan could indicate a common source, or could be an independent natural outcome of transparent semantics in creolization. The ultimate source of the distribution could be from a substrate language like Fongbe, the primary substrate language of the Surinamese creoles (Migge 1998), where a GOAL/BEN construction can be the GOAL/BEN SVC with V2 *ná* 'give', or a non-serial construction with the preposition *nú* 'to' (Lefebvre & Brousseau 2002: 446).

As for the transfer of SVCs into early Sranan, and the larger debate on the stabilization of Sranan and Saramaccan, the SVCs in Maroon Spirit Language offer no clue. The presence of SVCs in Maroon Spirit Language would indicate that the language of the Surinamese slaves brought to Jamaica was a stabilized creole, according to McWhorter (2000: 101) somewhat before 1665. An absence of SVCs in Maroon Spirit Language could be taken as in indication that this language had not stabilized by the date of the slaves' departure from Suriname in 1671. The presence of SVCs in Maroon Spirit Language as little changed from an expanded common pidgin might explain SVCs in all the English lexified creoles, which would entail the transfer of a non-essential item through the filter of an MIC pidgin, or indicate that this common language was in fact an MCS, namely a pidgin or creole used for community solidarity. But it wouldn't explain the SVCs in the French and Spanish/Portuguese lexified creoles. Smith (2002: 136) estimates the date of the creolization of Sranan at between 1665 and 1670. But Arends (2002: 124), on the basis of the 1718 Herlein document, argues that Sranan was not yet a stabilized creole by the early 18th century. Smith (2002: 135) argues that the creolization of Saramaccan was about 1680, involving the partial relexification of early Sranan with Portuguese. This process itself is considered to have been creolization (Smith 1999: 278), namely creolization of a stabilized creole. This would be not so much 'obliteration', or complete restructuring, as argued against by McWhorter (2000: 80), but rather a gentle recreolization. Considering that creoles can be recreolized with new contact without 'obliteration' means that SVCs could be transferred in a creolization process without upsetting already stabilized fea-

tures (Kramer 2005). Maroon Spirit Language, then, could have stabilized features also present in Sranan, but would not necessarily have had SVCs. Since similar SVCs occur in all the African substrate Atlantic creoles, it is most likely that SVCs transferred independently from common substrate Gbe languages, as there could be no common pidgin for creoles lexified by different European languages. Increased contact with speakers of Gbe languages, like Fongbe, would condition recreolization and allow for the late transfer of SVCs.

3. Conclusion

An indication that the posited pidgin predecessor to Saramaccan creole had no SVCs is the existence of non-serial alternatives for SVCs in the primary substrate language Fongbe, and in early and modern Saramaccan. As for input for the pidgin from the other important substrate language Kikongo, it is not a verb serializing language. The lexifier English, of course, is also not a verb serializing language, and has non-serial constructions that happen to match the word order of the substrate Fongbe non-serial alternatives, so there would be no need for SVCs in an English-Fongbe/Kikongo pidgin. Substrate Kwa SVCs developed over time and serve no greater function than the morphology they replaced, and like new morphology, the new SVCs complicated the grammar; they added both a syntactic and a semantic component. Like morphology, then, SVCs can be considered 'ornamental' and not a fundamental (McWhorter 2002: 4) feature of language. It is argued that nonessential features would not be found in creole languages because they are 'recent descendants of pidgins' (McWhorter 2002: 9), implying that such features would not transfer subsequent to the pidgin stage. Yet, SVCs transferred, they did not develop in the creole (McWhorter 1996). Their use in Saramaccan as stylistic variants is consistent with use in an expanded MCS stage creole, but not with use in a pidgin in the MIC stage.

The comparison of SVCs in Saramaccan and Fongbe shows that for every type of SVC the Saramaccan verbs have more verbal properties and are generally more semantically transparent, and are therefore less grammaticalized, than the corresponding verbs in Fongbe. This apparent reverse direction of grammaticalization of SVCs from Fongbe to Saramaccan is the result of transfer, with the predictable increase in verbal properties and semantic transparency in the creole over the source languages (McWhorter 2001). But this increase in verbal prototypicality in creoles would not necessarily be due to a pidgin filter as such (McWhorter 2002), but rather simply to the fact that there were speakers of non-serializing languages like Kikongo for whom the grammaticalization in the SVCs would be less intelligible, and therefore less essential. In the absence of English input in an MCS stage, speakers of Kikongo would be relatively more influential, able to block the grammaticalization of SVCs but not their transfer. The greater prototypicality of verbs in creole SVCs, then, does not necessarily imply transfer during a pidgin stage.

The comparison of SVCs in the late 18th century Saramaccan and modern Saramaccan shows that in the variation between SVCs and non-serial alternatives there is an increase in the use of SVCs in modern Saramaccan, even, as with GOAL/BEN depictions, to the extent of replacing the non-serial alternatives altogether. This increase could imply late incorporation (Bruyn 1996: 39), parallel to the time differences in the incorporation of Sranan *da* 'the' and *wan* 'a', discussed in Section 1. But the increase in frequency of SVCs in Saramaccan also means an increase in the use of a substrate feature, namely verb serialization. Yet, during this time there were very few African-born speakers of substrate languages in the Saramaccan community (Price 1983), as the majority of Saramaccan speakers would have been a few generations removed from the original substrate language speakers. It was the influence of these later generations of speakers, however, that increased the frequency of substrate verb serialization. Sankoff (1994: 314) argues that children acquire the constraints on variations from their parents accurately, but can vary the frequency; in Tok Pisin, children who are three or four generations removed from the original substrate speakers have increased the frequency of SVCs beyond the high frequency use of their parents, apparently due to the salience of SVCs by their high frequency use (p. 315).

Another factor in the increase in the frequency of salient substrate features, at least for Saramaccan, could be language identity (Le Page & Tabouret-Keller 1985). In 1762 the Saramaccans were successful in forcing the Dutch to grant them independence (Price 1983). In Belize, independence caused a redefining of language identity, resulting in an increase in the use of Belizian creole features (Le Page & Tabouret-Keller 1985). Self-ascribing a language identity "is always accompanied by linguistic symptoms", such as reification, that involves "some body of doctrine (grammar, lexicon, a literature) by means of which the system can be taught', and totemization, that uses language as one of the 'defining social properties of a group" (Le Page & Tabouret-Keller 1985: 236). If SVCs in Saramaccan were salient enough to cause an increase in the frequency of their use, they may also have been salient enough to be identified as a social property.

If salience can account for an increase in frequency for SVCs in Saramaccan, there is, then, the question of how SVCs would come to be salient. For the predecessor to the Surinamese creoles it has been argued here that SVCs likely transferred subsequent to an MIC stage, namely when communication with speakers of English could take place in an L2 English. The MIC would have no SVCs because there are none in English, there are similar non-serial constructions in the Gbe languages, and the speech community would include speakers of other African languages without SVCs. An MIC stage would allow only the most basic communication, while a subsequent and greatly expanded MCS stage would develop into a full language. The time of the English departure from Suriname, around 1680, was followed by a period of massive expansion of the slave population (Arends 2002: 121), the majority of whom were from the Slave Coast (Smith 2001: 56), and so spoke languages with SVCs. These people would be learning the developing creole language. At first contact with a language, Sankoff (1994) argues, the influence on that language by learners is limited. As they

become more fluent, there is greater influence. It is, however, at the point where that language becomes a usual medium of communication that the influence is greatest, because the native language can readily supply the means for communicating the expanded functions of a full language (Sankoff 1994). The proportion of African-born to Suriname-born slaves was extremely high due to the expansion of the plantation system, the low birth rate, and the replacement needed to make up for the high mortality (Arends 1995: 263). By the Sankoff model, then, substrate influence would have been considerable, and could have continued during the mass importation into the 18th century, increasing the use and salience, for example, of SVCs. The influence from verb serializing languages would have been considerable, as speakers from the Slave Coast made up about 40% of imported slaves from 1680–1699, and about 70% from 1700–1720 (Arends 1995: 243).

But implicit in the statement that creoles are 'recent descendants of pidgins' (McWhorter 2002: 9) is the claim that there is no late transfer of nonessential features, as part of the claim that such features will not be found in creoles unless they have time to develop. More directly, it is argued that for a nonessential feature not found in pidgins 'there is a very strong tendency for creoles to eschew that feature even if said feature is well-represented in the source language and/or expressed analytically' (McWhorter 2002: 33). There are nonessential features in creole languages, though, such as tense and aspect marking (p. 34), explained as creolization operating in degrees by selecting certain features, such as tense and aspect marking, before others. Tense and aspect would seem more essential than SVCs, if for no other reason than, unlike SVCs, most languages have tense and aspect marking. As for ease of transfer, SVCs complicate the creole grammar with syntax that is at odds with the basic creole SVO word order, and with special semantics. Now in Saramaccan both the TMA system and SVCs match Fongbe, and are most likely the result of transfer. It would seem that if features were selected by degrees, in this case selected for transfer, that the TMA system would transfer first. Since pidgins are characterized as using adverbs instead of TMA systems (Smith 2001: 62), TMA transfer must have occurred in the creole or MCS stage. Baker (1995: 6–7) shows that the preverbal past marker and the preverbal future marker occur in Sranan by 1718, when there was a zero copula and no definite article, but the preverbal progressive marker, a 'creole' characteristic (p. 14) occurs first in 1765, as there is no context for it in the 1718 text. These are well after the proposed 1665 date of creolization of Sranan (Smith 2002: 36). As for SVCs, there are none in the 1718 text, although *take jou oudy* 'sends you greetings' (Arends & Perl 1995: 74) in later texts might also have been a 'talk-give' SVC. A number of types of SVC are missing from the 1770 and the lengthy 1765 texts that occur in Schumann 1783 (Arends & Perl 1995: 56). This may be an indication that the frequency of the use of SVCs had been increasing, implicating a somewhat recent transfer.

As the notion of markedness relates to transfer, a marked syntactic device would be considered nonessential, since 'creoles represent unmarked syntactic parameters' (McWhorter 2002: 39), and therefore would not transfer. In this sense, syntactic constructions, as devices for handling essential items (McWhorter 1996: 155), do need

to be considered either essential or nonessential in terms of markedness. Although there is no particular agreement on markedness of SVCs, SVCs in Saramaccan can arguably be considered marked because of their semantic constraints. In both Fongbe and Saramaccan there are constraints on SVCs that do not apply to the non-serial alternatives, such as constraints against ACTIVITY depictions. Non-serial alternatives then would be unmarked compared to SVCs because they can occur where SVCs occur, and where they can't, such as in ACTIVITY depictions. SVCs could also be considered marked in that in both Fongbe and Saramaccan the serial item verb, usually V2, has fewer verbal qualities than the matrix verb, such as no independent tense marking. Two equivalent alternative single-verb clauses would have independent tense marking. As for discourse functions, the shared object SVC in particular, in 2.1.6, functions to introduce or reactivate discourse participants, as it rarely occurs with pronominal objects. The non-serial alternatives, of course, occur with pronominal objects and named objects, so they may also function to introduce or reactivate participants. Lambrecht (1995: 17) states a general rule that 'given a pair of allosentences, one member is pragmatically unmarked if it serves two discourse functions while the other member serves only one of them. While the marked member is positively specified for some pragmatic feature, the unmarked member is neutral with respect to this feature.' It would seem, then, that the various SVCs could be considered marked for various features, and yet have transferred into a number of creole languages.

The McWhorter model of pidginization, and its relation to the concept of essential features, could provide an explanation for the late transfer of features into Saramaccan; rather than restricting input to essential features, pidginization should be seen as requiring the input of essential features. There are certainly minimal features necessary for language, necessarily forged in pidginization in the absence of common features, the winning feature being selected for reasons such as simplicity, salience, or social dominance. On the other hand, absence of common input features can explain the general absence of the features considered nonessential in pidgins, since the randomness of grammaticalization and reanalysis would rarely line up common nonessential features, especially in an MIC. For essential features, then, the motivation to create a basic communication system determines the inclusion of essential features, even if there are no common features, but there is no such motivation for nonessential features. But in an MCS like the predecessor to Saramaccan, there is the likelihood of increasing the inventory of common features because many of the remaining contact languages are similar, or of increasing dominant language features due to solidarity motivating familiarity with the dominant language. The closed set of minimal features necessary for natural language would define the minimum, not the maximum inventory of features in a creole, the inventory of 'nonessential' items being determined by sociohistorical and typological factors. Creolization would occur whenever contact resulted in solidarity and a redefining of language identity, something that probably happened often in the history of Saramaccan.

In the predecessors to Saramaccan, namely an MIC pidgin and an MCS pidgin or creole expanding to early Sranan, the SVO word order is generally common to English

and Fongbe, the major input languages. But the English prenominal directionality for attributive adjectives prevailed over the Fongbe postnominal, perhaps because in the MIC stage the English were socially dominant; Hudson (1980: 62) argues that in a contact situation those who have the least to benefit from learning a contact language, here it was the English, will compromise least linguistically. Now in the MCS stage, speakers of Gbe languages were obviously dominant, and the English were by definition not present. Among the Gbe languages the basic syntax and semantics of SVCs is similar, so there would be no obstacle to their incorporation into the creole grammar. The dominance by Gbe speakers can be partly explained by their increasing proportion of the population (Smith 2002; Arends 1995), but also by the fact that the Gbe languages themselves have creole characteristics in that they are mainly isolating, are basically SVO, and have TMA systems. The complex morphology of Kikongo, whose speakers were the other main substrate population, would have been lost, while the Gbe syntax that was retained basically matched English. But in the complex history of Saramaccan it seems reasonable that the conditions for creolization could recur and that Gbe would not always dominate; the Portuguese element in Saramaccan, for example, is considered by many to have entered the language through creolization (Smith 1999: 278), involving some form of Portuguese and an existing early Sranan.

In the complex history of Saramaccan, each type of SVC no doubt has its own history as well. Some were probably early, others appear to have become more common over time, and others do not appear in the early texts. But they all have one thing in common; they are marked, as argued above, by being constrained, compared to the non-serial alternatives. Like the SVCs, the marked characteristics, namely their constraints, also transferred, so that the SVCs were most likely marked from their very first occurrence. Partly because of their specialized use, then, they most likely transferred after the non-serial alternatives. And because they could be considered nonessential, they most likely transferred after the MIC stage, in a contact situation characterized by solidarity and redefined language identity.

References

Aboikoni, L. & Glock, N. (1997). *Di duumi u Gaama Aboikoni*. Paramaribo: Summer Institute of Linguistics.

Alleyne, M. (1980). *Comparative Afro-American: An historical-comparative study of English-based Afro-American dialects of the new world.* Karoma Publishers MI: Ann Arbor.

Alleyne, M. (2000). Opposite processes in creolization. In I. Neumann-Holzschuh & E. Schneider (Eds.), *Degrees of Restructuring in Creole Languages* (pp. 125–133). Amsterdam:John Benjamins.

Arends, J. (1989). Syntactic Developments in Sranan; Creolization as a Gradual Process. Ph.D dissertation, Katholieke Univertsiteit Nijmegen.

Arends, J. (1995). Demographic factors in the formation of Sranan. In J. Arends (Ed.), *The Early Stages of Creolization* (pp. 233–285). Amsterdam: John Benjamins.

Arends, J. (2002). The history of the Surinamese creoles I. A sociohistorical survey. In E. Carlin & J. Arends (Eds.), *Atlas of the Languages of Suriname* (pp. 115–130). Leiden: KITLV Press.

Arends, J. & Perl, M. (1995). *Early Suriname Creole Texts*. Frankfurt: Vervuert.

Baker, P. (1995). Some developmental inferences from the historical studies of pidgins and creoles. In J. Arends (Ed.), *The Early Stages of Creolization* (pp. 1–24). Amsterdam: John Benjamins.

Baker, P. (2000). Theories of creolization and the degree and nature of restructuring. In I. Neumann-Holzschuh & E. W. Schneider (Eds.), *Degrees of Restructuring in Creole Languages* (pp. 41–64). Amsterdam: John Benjamins.

Bilby, K. M. (1981). The Kromanti dance of the windward maroons of Jamaica. *Nieuwe West-Indische Gids, 55*, 52–101.

Bilby, K. M. (1983). How the older heads talk: A Jamaican Maroon Spirit possession language and its relationship to the creoles of Suriname and Sierra Leone. *Nieuwe West-Indische Gids, 57*, 37–88.

Bilby, K. M. (1992). Further observations on the Jamaican Maroon Spirit Language. Paper presented at the 1992 Annual Meeting of the Society for Pidgin and Creole Linguistics, Philadelphia.

Bruyn, A. (1996). On identifying instances of grammaticalization in creole languages. In P. Baker & A. Syea (Eds.), *Changing Meanings, Changing Functions* [Westminster Creolistics Series 2] (pp. 29–46). London: University of Westminster Press.

Byrne, F. (1987). *Grammatical Relations in a Radical Creole: Verb complementation in Saramaccan.* Amsterdam: John Benjamins.

Crowley, T. (2002). *Serial Verbs in Oceanic: A descriptive typology*. New York NY: OUP.

da Cruz, M. (1994). Les constructions sérielles du fongbè. Approches sémantique et syntaxique. Ph.D dissertation, Université du Québec à Montréal.

Derbyshire, D. C. (1979). *Hixkaryana. Lingua Descriptive Studies* V. 1. Amsterdam: North Holland.

Dowty, D. (1979). *Word Meaning and Montague Grammar: The semantics of verbs and times in generative semantics and in Montague's PDQ.* Dordrecht: Reidel.

Du Bois, J. (1987). The discourse basis of ergativity. *Language, 63*, 805–855.

Duranti, A. (1990). Politics and grammar: Agency in Samoan discourse. *American Anthropologist, 17*, 36–56.

Durie, M. (1997). Grammatical structures in verb serialization. In A. Alsina, J. Bresnan, & P. Sells (Eds.), *Complex Predicates* [CSLI Publications 64] (pp. 289–354). Stanford CA: CSLI.

Givón, T. (1991). Some substantive issues concerning verb serialization: Grammatical versus cognitive packaging. In C. Lefebvre (Ed.), *Serial verbs: Grammatical, comparative and cognitive approaches* (pp. 137–184). Amsterdam: John Benjamins.

Glock, N. (1972). Clause and sentence in Saramaccan. *Journal of African Languages, 11*, 45–61.

Hale, K. & Keyser, J. (1987). A view from the middle. *MIT Lexicon Project Working Papers 10.* Cambridge MA: Center for Cognitive Science, MIT.

Hudson, R. A. (1980). *Sociolinguistics*. Cambridge: CUP.

Hyman, L. (2000). How to become a Kwa verb. Paper presented at the *Symposium on Areal Typology of West African Languages*, Leipzig, September 1–2, 2000.

Kihm, A. (1989). Lexical conflation as a basis for relexification. *Canadian Journal of Linguistics, 34*, 351–376.

Kramer, M. (2002). Substrate Transfer in Saramaccan Creole. Ph.D dissertation, University of California, Berkeley.

Kramer, M. (2005). High tone spread in Saramaccan serial verb constructions. *Journal of Portuguese Linguistics, 3*, 31–53.

Kramer, M. (2006). Inalienable possession in Saramaccan, ms.

Kramp, A. A. (1983). Neger-Englisches Wörterbuch. Early Creole Lexicography: A study of C. L. Schumann's manuscript dictionary of Sranan. Ph.D dissertation, Leiden University, 44–305.

Labov, W. (1990). On the adequacy of natural languages: The development of tense. In J. Singler (Ed.), *Pidgin and Creole Tense-mood-aspect Systems* (pp. 1–58). Amsterdam: John Benjamins.

Lambrecht, K. (1995). *Information Structure and Sentence Form* [Cambridge studies in linguistics 71]. Cambridge: CUP.

Lefebvre, C. & Brousseau, A.-M. (2002). *A Grammar of Fongbe*. Berlin: Mouton de Gruyter.

Le Page, R. & Tabouret-Keller, A. (1985). *Acts of Identity*. Cambridge: CUP.

Migge, B. (1998). Substrate Influence in the Formation of the Surinamese Plantation Creole: A consideration of sociohistorical data and the linguistic data from Ndyuka and Gbe. Ph.D dissertation, Ohio State University, Columbus.

McWhorter, J. H. (1996). *Towards a New Model of Creole Genesis*. New York NY: Peter Lang.

McWhorter, J. H. (2000). *The Missing Spanish Creoles: Recovering the birth of plantation contact languages*. Berkeley CA: University of California Press.

McWhorter, J. H. (2001). The world's simplest grammars are creole grammars. *Linguistic Typology, 5*, 125–156.

McWhorter, J. H. (2002). The rest of the story: Restoring pidginization to creole genesis theory. *Journal of Pidgin and Creole Languages, 17*, 1–48.

McWhorter, J. H. (2004). Saramaccan and Haitian as young grammars: The pitfalls of syntactocentrism in Creole genesis research. *Journal of Pidgin and Creole Languages, 19*, 77–137.

Ndayiragije, J. (1992). Clivage du prédicat en Fongbè. In A. Kihm & C. Lefebvre (Eds.), *Aspects de la grammaire du Fongbè* (pp. 61–86). Paris: Peeters.

Nichols, J. (1992). *Linguistic Diversity in Space and Time*. Chicago IL: University of Chicago Press.

Plag, I. (1993). *Sentential Complementation in Sranan*. Tübingen: Niemeyer.

Price, R. (1983). *First-time. The historical version of an Afro-American people*. Baltimore MD: Johns Hopkins University Press.

Rountree, C. (1972). Saramaccan tone in relation to intonation and grammar. *Lingua, 29*, 308–325.

Sankoff, G. (1994). An historical and evolutionary approach to variation in the Tok Pisin verb phrase. In K. Beals et al. (Eds.), *Papers from the 30th Regional Meeting of the Chicago Linguistic Society*. V 2: *The parasession on variation in linguistic theory* (pp. 293–320). Chicago IL: Chicago Linguistic Society.

Schuchardt, H. (1914). *Die Sprache der Saramakkaneger in Surinam* [Verhandelingen der Koninklijke Akademie van Wetenschappen te Amsterdam, Afdeeling Letterkunde. Nieuwe reeks. Deel XIV no. 6]. Amsterdam: Johannes Müller.

Schumann, C. L. (1778). Saramaccanisch Deutsches Wörter-Buch. In Schuchardt (1914) (Ed., pp. 46–116).

Sebba, M. (1987). *The Syntax of Serial Verbs: An investigation into serialization in Sranan and other languages* [Creole Language Library 2]. Amsterdam: John Benjamins.

Smith, N. (1999). Pernambuco to Surinam 1654–1665? The Jewish slave controversy. In M. Huber & M. Parkvall (Eds.), *Spreading the Word* [Westminster Creolistics Series 6] (pp. 251–298). London: University of Westminster Press.

Smith, N. (2001). Voodoo chile: Differential substrate effects in Saramaccan and Haitian. In N. Smith & T. Veenstra (Eds.), *Creolization and Contact* [Creole Language Library 23] (pp. 43–80). Amsterdam: John Benjamins.

Smith, N. (2002). The history of the Surinamese creoles II: Origin and differentiation. In E. Carlin & J. Arends (Eds.), *Atlas of the Languages of Suriname* (pp. 131–152). Leiden: KITLV Press.

Summer Institute of Linguistics (1991). *Gadu Buku* (*The New Testament in Saramaccan*). Paramaribo: SIL.

Thomason, S. G. & Kaufman, T. (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley CA: University of California Press.

Veenstra, T. (1996). *Serial Verbs in Saramaccan: Predication and creole genesis*. The Hague: HIL.

Wietz, Br. (1805). Die Apostel-Geschichte in die Saramakka-Neger-Sprache. In Schuchardt 1914.

Whinnom, K. (1971). Linguistic hybridization and the 'special case' of pidgins and creoles. In D. Hymes (Ed.), *Pidginization and Creolization of Languages* (pp. 91–116). Oxford: Alden Press.

Index

A

AAVE 203ff., 225f., 228, 233, 239, 241, 284
accent 12, 15, 23ff., 144, *see also* primary accent, secondary accent
acrolect/acrolectal 245, 256f., 258, 263f.
adjectival predicates 205
affixation 63ff., 117, 142f., 316, 323f.
affix-hopping 208
African languages 15, 17, 86, 279, 286, 338, 358, 366
Afrikaans 284, 326
Afro-American creoles 363
Afro-Hispanic 265ff.
Afro-Iberian pidgin 263, 265, 271
Afro-Portuguese pidgin 265ff., 272
age 68, 70f., 226, 250f., 300f., 306, 308f.
agentive *be* 203ff.
agglutination 319ff.
Aja, *see* Gbe
Alsea 326
alienable/inalienable possession 341
allosentences 340, 368
Amerindian 16, 111ff.
Amharic 326
apparent time 294, 300, 307, 309
Arawakan 16
Assamese 326
aspectual 'finish' 30, 90, 93f. 97, 100, 102, 345ff.
Atlantic English creoles 363
Austronesian 10
auxiliary 45, 49, 54, 104, 105, 142f., 166, 169, 171, 191, 204, 206, 208f., 211, 215, 218, 220, 227f., 230f., 252, 254, 310
aux raising 119
aux inversion 122, 230
Azerbaijani 326

B

Bahamian Creole English 225ff.
Bantu languages 30
Barbadian Creole 89
basic lexicon 322, 329
Basic Variety 67
basilect/basilectal 245f., 256f., 258, 285, 363f.
basilectalization 86ff.
Basque 116, 326
Belarusian 184, 189, 191
bilingualism/bilingual 153f., 156, 161, 176, 188, 243, 257f., 295, 298, 299, 321, 326, 338
bilingual interference 185
Black speech 264, 271, 274, 281, 283, 284, 286
borrowing 103, 326ff.
bozal Spanish 280
bozales 270
Brazilian Portuguese 284, 285, 286
broken transmission hypothesis 321f., 327, 330

C

calquing 176, 179
Canadian French 92, 320
Cancioneiro Geral 271
Cantonese 326
Cape Verdean Creole 276, 284, 286
Caribbean creoles 54, 105, 193, 222, 225f., 228, 239, 241, 246, 256, 257, 294f., 338, 363
Caribbean plantation pidgin 338
Caribbean Spanish 286
case marking 145ff.
change from above 264, 294, 300, 307, 308
'chat' mode 229ff.
Chinese 89, 264
Chinook Jargon 111ff.
Chinookan 112, 125
clitic 130, 164
code-switching 252, 257
cofradías 267
colonial German 186
completive marker 30, 92f., 97, 99, 100ff., 230
complicating action (CA) clauses 229f., 231, 235f.
consonant cluster reduction 225
conversion/converted language 138, 160, 178f.
continuity hypothesis 321
Coos 326
copula 31, 98, 100, 103, 191, 208, 211, 214, 218, 221, 227, 230f., 244, 246, 255ff., 277, 293, 310, 328
creole archetype 176, 179
creole genesis 30, 57, 65ff., 85ff., 94, 104f., 112, 131, 160, 311, 329, 337f.
creole transparency 362
creolization 61ff., 88, 160, 175f., 178f., 243, 281f., 311, 315ff., 338, 364, 367
Cuban *bozal* speech 284
CV syllable 13f., 269, 286

D

decreolization 221, 245f., 258f., 311, 321, 327
deletion 218ff., 265, 271f., 277, 284f.
demography/demographic 91, 106, 154, 294
deontic 34ff., 48ff., 172f.
Devon English 217

diachrony/diachronic 31, 63, 78ff., 150, 156, 293f., 339, 346
diglossia/diglossic 163, 178, 249, 338
discontinuity 63, 65, 80
ditransitives 356f.
domain contrast 144, 153
do-support 119, 204f., 208ff., 218ff., 227, 231
DP predicates 206
durative 102, 160, 233, 303, 306f.
Dutch 16, 32f., 47f., 56f., 71, 78, 97ff., 103, 105, 295ff., 307, 310f., 328

E
economic change 299
emphatic pronoun 15, 25
English 13, 16, 23, 32, 46, 48, 50f., 55f., 64f., 66, 68ff., 88, 97ff., 101, 111, 118, 122, 161, 203ff., 233, 243ff., 292ff., 320, 322, 325, 327, 338f., 344, 352, 363, 365, 368f.
English-Fongbe pidgin 252, 365
epenthetic vowel 16, 264, 269
epistemic possibility 34f., 39, 42f., 47f., 172
Esperanto 189
Estonian 326
expanded CP 116f., 130

F
F_0 18ff.
Flemish 119f.
folktales 228, 234ff.
Fongbe 30f., 86, 95, 101, 338ff.
future marker 30, 42, 47f., 50, 52, 91ff., 104f., 142, 144, 147f., 164, 168f., 175f. 252, 367
future potential marker 38, 103
French 30ff., 86, 88, 91ff., 111, 220, 269, 298, 316ff., 363f.
front rounded vowels 316f.

G
Gen, *see* Gbe
gender agreement 270, 276, 283f., 316, 324, 326
Gbe 30ff., 64f., 70, 72, 94f., 99ff., 337ff.
grammaticalization 78, 86, 94, 105, 311, 339, 342f., 365, 368
grammatical competition 139
Guinea-Bissau Portuguese Creole 328
Guyanese Creole 10, 239, 256

H
habla de negro 266, 275, 277
habitual *be* 203ff.
habitual 32, 100f., 103, 105, 228f., 237ff., 252f., 302, 303, 306
Hausa 249
Haitian Creole 30f., 86ff., 120, 315ff.
Hawai'i Creole English 105
Hawai'i Pidgin English 89
head-initial 68, 71f., 77
Hebrew 187ff.
Helvécia 270
historical present 232
Hixkaryana 341

I
identity (social) 50, 161, 184, 327, 366, 369
ideophones 16, 19
Igbo 249
Ilocano 324, 326
immigration/migration 177, 184, 187, 296, 299, 310
imperfective 36, 48, 90, 96f., 99, 100, 102ff., 167, 252, 293, 302f., 305f., 308
inflectional morphology 90, 338
innovators 61ff.
input 31, 72, 85, 87, 89, 99, 105f., 300, 329, 339, 365
Input Correspondence Rule 62, 74, 75f., 77
intake 87
interlanguage 69, 72, 87, 89, 96, 106
internal factors 310f., 341
intervocalic *d* 279
intonational language 17
Irish English 221f.
irrealis 30, 52, 160, 175, 352
Isle de France Creole 30
iterative 102, 303, 306f., 312

J
Jamaican Creole 245f., 255f., 327, 263f.

K
Karaite 186
Kiev-Polessian 184, 189, 190f., 195
Kikongo 99, 104, 338f., 365, 368f.
koiné 156
Kongo 17
Korean 326
Krio 10
Kwa 30, 90, 95, 337f.

L
Latin 116, 194, 196, 321, 325f.
lexicon 9ff., 70, 74f., 138, 160, 162, 184ff., 249, 321f., 323ff.
language bioprogram hypothesis 30, 66, 85, 104, 175
L1 acquisition 56, 61, 65, 70ff., 79, 85, 95
L2 acquisition 57, 63, 65, 67ff., 79, 85ff., 150, 153f., 279, 286, 315, 320f.
learning strategy 93f.
lexicography 192ff.
língua de preto 266, 283
lingua de guiné 283
lingua franca 111, 161f., 243f.
literatura de cordel 282
Lithuanian 190, 326
Louisiana French 320, 328
Low German 326

M
Malay 30, 135ff., 160ff.
mapping rules/principles 73ff.
Mari 326
markedness 367f., 369
maroon creoles 31ff.
Maroon Spirit Language 363ff.
Mauritian French Creole 328
Maxi, *see* Gbe
maximal projection 62, 75, 77, 146
Ma'ya 10
medium for community solidarity 338, 365
medium for interethnic communication 338, 365
Megleno-Romanian 326
Melanesian pidgin 105
mesolect/mesolectal 245f., 256ff.
Mexico City Spanish 310

Middle High German 188f., 193, 195
modal verbs 31, 33, 44, 55, 228, 253f.
modality 34ff., 104, 172, 252f.
morphological compounding 66f., 71, 73, 78ff., 323
morphological change 78, 323f., 326
multilingualism, *see* bilingualism
multivariate analysis 294, 305

N
-ndo 291ff.
nativization 61ff., 154
Ndunga-le 326
Ndyuka 31ff., 62ff.
negation 112ff., 140ff., 204f., 208ff., 212, 215, 218ff., 252, 256
negative concord 117ff.
negative indefinites 118, 123, 129f.
negative operators 122ff.
negative possibility 39ff.
negative physical ability marker 49
negative quantifiers 118, 123ff.
NegP 116ff., 211, 213, 220
New Zealand English 209, 222
Ngandi 326
Nigerian Pidgin English 243ff.
nominal subject 113ff.
nominalizations 68, 126, 144ff.
non-serial alternatives 342ff.

O
object raising 138
obligation 34, 47, 50ff., 104, 172
oral narratives 225ff.
order of acquisition (tense/aspect) 90
overt agreement constraint 208
overt inflection constraint 207ff., 218

P
Pamaka 31ff.
pan-Caribbean creole language 286
Papiamentu 10, 276, 284, 286, 291ff., 328
paradigmatic contrast 14f.
partial restructuring 281f., 284, 286
past marking 90f., 93, 96f., 99ff., 160, 165, 166f., 225ff., 232f., 234, 239f., 246, 252, 256, 267
perfective 100f., 169ff., 232, 252
Persian 326
phonetics 9ff.
phonological phrases 10f., 21ff., 27
phonological prominence 13f.
phonology 9ff. 73, 75f., 183, 185, 198, 324
phonosyntactic 10f.
pidgin 72, 85f., 88f., 98f., 111ff., 162, 243ff., 280, 315ff., 321, 337ff., 364, 365
pidgin filter 337f., 339, 342, 365
pidginization 72, 315, 368
pitch 17, 20f., 24
pitch accent 9, 12ff.
pitch trace 15
plateauing 10ff., 18
plural marking 268, 270, 273ff.
Portuguese 16, 33, 46f., 49f., 57, 89, 97, 99, 138, 264f., 267f., 272, 275f., 279, 282, 363f., 369
Portuguese-based creoles 30, 98, 138, 268f., 284, 286, 328, 364
post-vocalic /r/ 316f.
predicate, semantics of 303f.
pre-nominal relative clauses 152f.
preverbal markers 31, 37, 86, 95f., 100, 102, 104f., 164, 173, 227, 231, 252f., 255f., 367
primary accent 22f.
progressive 36, 91ff., 166, 171f., 176, 206, 210, 216, 219f., 221, 252, 257, 291ff., 367
pronominal object 101, 344, 352, 368
pronomial subject 113
prosody/prosodic 10, 17, 112, 114f., 144

Q
Quechua 326

R
reanalysis 49, 57, 86, 88, 91, 94, 99, 106, 138, 142f., 147f., 149f., 156, 221, 368
recreolization 364f.
reference time 169, 228, 232f., 237, 239
regulators 61, 63
Relativized Minimality 212, 216, 220
relexification 30f., 57, 66, 86, 154, 183ff., 344, 364
repeated action 303, 305f.
Representational Modularity 73
Résigaro 326
restructuring 61, 87f., 95, 105f., 281, 364, *see also* partial restructuring
Reunion creole 89
Reunion French 284

S
Saamaka 31ff.
Saami 326
salience 93, 310, 323, 366, 368
Salishan 112
Samaná English 217
Samoan 341
São Tomense 284, 286
Saramaccan 9ff., 62ff., 88, 98, 337ff.
sayagués 272
secondary accent 22ff.
segmental phonology 27
semi-creole 320, 284
sentence-initial negation 125
sentential negation 111ff.
Sephardic Spanish 264
Serbian 117, 326
serial verbs 337ff.
sex 226, 300f.
Sinhala 136ff., 161f., 166, 168f., 175f., 178
Siuslaw 326
social class 234, 300f.
social prestige 307f.
sociolinguistic constraints 294, 308
sociolinguistic interviews 226, 234, 294, 302
sociolinguistic change 88
sociolinguistic variable 307
Sorbian 183ff.
South Seas Pidgin 89
Spanish 68, 264f., 267, 269, 271f., 274, 277ff., 292ff., 323f., 327f., 363

Spanish-based creoles 269
Sranan 31ff., 62, 78, 86, 88f., 94, 96ff., 297, 311, 328, 339, 342, 344, 346f., 350, 357ff.
Sri Lanka English 175
Sri Lanka Portuguese 162, 164, 178
Sri Lanka Malay 136ff., 159ff.
Stammbaum theory 325
stative 100f., 126, 231, 233, 238, 252, 303ff., 313
stress 16, 23, 189f.
stress accent 13, 23
stylistic variation 341f.
subjunctive 30, 35, 53f., 94, 204, 207, 209ff., 219, 253
substrate 30, 31, 47f., 49, 56f., 62, 64f., 68, 80, 85f., 89, 91, 93ff., 131, 153, 160, 162f., 176, 179, 258, 269, 377ff.
superstrate 30, 31, 47, 56f., 62, 64f., 68, 80, 85f., 87, 88f., 91, 93ff., 163, 185, 187ff., 258, 329
surface string matching 310
Surinamese creoles 16, 29ff., 62f., 67, 72, 79, 99, 101f., 328, 337ff., 363
syntagmatic contrast 14
synthetic compounds 62ff.
syntactic change 264
syntactic constraints 131
syntactic object 75f., 123
syntactic primitives 73
Swadesh list 325f., 330
Swedish 325f.

T
Tagalog 326
Tajuki 324, 326
Tamil 136ff., 160ff.
target language 86f., 88, 105
Tat 326
tertiary hybridization 339
TMA marking 29ff., 85ff., 143f., 160ff., 240, 293, 303, 346, 367, 369
Tok Pisin 246, 258, 366
Tongan 341
tone 9ff., 65, 77, 101, 344
transfer 17, 106, 112, 286, 326, 337ff.
Trinidadian Creole 239

U
Ukrainian 184ff.
urban/rural 246, 295, 300f.

V
Varbrul 239, 305ff.
variability/variation 39, 78, 116, 137, 139, 150, 168, 225ff., 243ff., 291ff., 337ff.
Venezuelan Spanish 309
verbhood tests 36
verb movement 211ff.
verb-object order 64, 69f.
villancicos 284
VO/OV variation 138f.
vP recursion 217
voiceless stop 26
vowel harmony 18

W
Waci, *see* Gbe
Welsh 326
WH movement 150
WH criterion 119

X
Xwela, *see* Gbe
Xwla, *see* Gbe

Y
Yiddish 183ff., 326
Yoruba 249, 257
Yupik 326

Z
zero affixation 77, 80

In the series *Creole Language Library* the following titles have been published thus far or are scheduled for publication:

31 **BAPTISTA, Marlyse and Jacqueline GUÉRON (eds.):** Bare NP's in Creole Languages. *Expected August 2007*

30 **SIEGEL, Jeff, John LYNCH and Diana EADES (eds.):** Language Description, History and Development. Linguistic indulgence in memory of Terry Crowley. xvi, 512 pp. *Expected April 2007*

29 **DEUMERT, Ana and Stephanie DURRLEMAN (eds.):** Structure and Variation in Language Contact. 2006. viii, 376 pp.

28 **CLEMENTS, J. Clancy, Thomas A. KLINGLER, Deborah PISTON-HATLEN and Kevin J. ROTTET (eds.):** History, Society and Variation. In honor of Albert Valdman. 2006. vi, 304 pp.

27 **ESCURE, Geneviève and Armin SCHWEGLER (eds.):** Creoles, Contact, and Language Change. Linguistic and social implications. 2004. x, 355 pp.

26 **MOUS, Maarten:** The Making of a Mixed Language. The case of Ma'a/Mbugu. 2003. xx, 322 pp.

25 **MIGGE, Bettina:** Creole Formation as Language Contact. The case of the Suriname Creoles. 2003. xii, 151 pp.

24 **MÜHLEISEN, Susanne:** Creole Discourse. Exploring prestige formation and change across Caribbean English-lexicon Creoles. 2002. xiv, 332 pp.

23 **SMITH, Norval and Tonjes VEENSTRA (eds.):** Creolization and Contact. 2001. vi, 323 pp.

22 **NEUMANN-HOLZSCHUH, Ingrid and Edgar W. SCHNEIDER (eds.):** Degrees of Restructuring in Creole Languages. 2001. iv, 492 pp.

21 **McWHORTER, John (ed.):** Language Change and Language Contact in Pidgins and Creoles. 2000. viii, 503 pp.

20 **RICKFORD, John R. and Suzanne ROMAINE (eds.):** Creole Genesis, Attitudes and Discourse. Studies celebrating Charlene J. Sato. 1999. viii, 418 pp.

19 **SPEARS, Arthur K. and Donald WINFORD (eds.):** The Structure and Status of Pidgins and Creoles. Including selected papers from meetings of the Society for Pidgin and Creole linguistics. 1997. viii, 461 pp.

18 **ESCURE, Geneviève:** Creole and Dialect Continua. Standard acquisition processes in Belize and China (PRC). 1997. x, 307 pp.

17 **THOMASON, Sarah G. (ed.):** Contact Languages. A wider perspective. 1997. xiii, 506 pp.

16 **CLEMENTS, J. Clancy:** The Genesis of a Language. The formation and development of Korlai Portuguese. 1996. xviii, 282 pp.

15 **ARENDS, Jacques, Pieter MUYSKEN and Norval SMITH (eds.):** Pidgins and Creoles. An introduction. 1994. xv, 412 pp.

14 **KIHM, Alain:** Kriyol Syntax. The Portuguese-based Creole language of Guinea-Bissau. 1994. xii, 310 pp.

13 **ARENDS, Jacques (ed.):** The Early Stages of Creolization. 1996. xvi, 297 pp.

12 **BYRNE, Francis and Donald WINFORD (eds.):** Focus and Grammatical Relations in Creole Languages. Papers from the University of Chicago Conference on Focus and Grammatical Relations in Creole Languages. 1993. xvi, 329 pp.

11 **BYRNE, Francis and John HOLM (eds.):** Atlantic Meets Pacific. A global view of pidginization and creolization. 1992. ix, 465 pp.

10 **WINFORD, Donald:** Predication in Caribbean English Creoles. 1993. viii, 419 pp.

9 **BYRNE, Francis and Thom HUEBNER (eds.):** Development and Structures of Creole Languages. Essays in honor of Derek Bickerton. 1991. x, 222 pp.

8 **BAILEY, Guy, Natalie MAYNOR and Patricia CUKOR-AVILA (eds.):** The Emergence of Black English. Text and commentary. 1991. x, 352 pp.

7 **FABIAN, Johannes (ed.):** History from Below. The "Vocabulary of Elisabethville" by André Yav: Text, Translations and Interpretive Essay. With the assistance of Kalundi Mango. With linguistic notes by W. Schicho. 1990. vii, 236 pp.

6 **SINGLER, John Victor (ed.):** Pidgin and Creole Tense/Mood/Aspect Systems. 1990. xvi, 240 pp.

5 **JACKSON, Kenneth David:** Sing Without Shame. Oral traditions in Indo-Portuguese Creole verse. 1990. xxiv, 257 pp.

4 **LIPSKI, John M.:** The Speech of the Negros Congos in Panama. 1989. vii, 159 pp.

3 **BYRNE, Francis:** Grammatical Relations in a Radical Creole. Verb Complementation in Saramaccan. With a foreword by Derek Bickerton. 1987. xiv, 293 pp.

2 **SEBBA, Mark:** The Syntax of Serial Verbs. An investigation into serialisation in Sranan and other languages. 1987. xv, 218 pp.

1 **MUYSKEN, Pieter and Norval SMITH (eds.):** Substrata versus Universals in Creole Genesis. Papers from the Amsterdam Creole Workshop, April 1985. 1986. vii, 311 pp.