

New research on circum-Caribbean creoles and language contact

Edited by

Angela Bartens

Peter Slomanson

Kristoffer Friis Bøegh

Studies in Caribbean Languages 6



Studies in Caribbean Languages

Chief Editor: Joseph T. Farquharson

In this series:

1. Irvine-Sobers, G. Alison. The acrolect in Jamaica: The architecture of phonological variation.
2. Forbes-Barnett, Marsha. Dual aspectual forms and event structure in Caribbean English Creoles.
3. Sherriah, André Ché. A tale of two dialect regions: Sranan's 17th-century English input.
4. Delgado, Sally. Ship English: Sailors' speech in the early colonial Caribbean.
5. Forrester, Clive (ed.). Intersections of language rights and social justice in the Caribbean context.
6. Angela Bartens, Peter Slomanson & Kristoffer Friis Bøegh (eds.). New research on circum-Caribbean creoles and language contact.

ISSN (print): 2627-1893
ISSN (electronic): 2627-1834

New research on circum-Caribbean creoles and language contact

Edited by

Angela Bartens

Peter Slomanson

Kristoffer Friis Bøegh



Angela Bartens, Peter Slomanson & Kristoffer Friis Bøegh (eds.). 2024. *New research on circum-Caribbean creoles and language contact* (Studies in Caribbean Languages 6). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/430>

© 2024, the authors

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/> 

ISBN: 978-3-96110-496-3 (Digital)

978-3-98554-125-6 (Hardcover)

ISSN (print): 2627-1893

ISSN (electronic): 2627-1834

DOI: 10.5281/zenodo.14281420

Source code available from www.github.com/langsci/430

Errata: paperhive.org/documents/remote?type=langsci&id=430

Cover and concept of design: Ulrike Harbort

Typesetting: Sebastian Nordhoff

Proofreading: Amir Ghorbanpour, Christopher Green, Diana Lewis, Elliott Pearl, Harold Somers, Jean Nitzke, Jeroen van de Weijer, Laura Arnold, Mary Ann Walter, Piero Visconte, Tom Bossuyt

Fonts: Libertinus, Arimo, DejaVu Sans Mono

Typesetting software: X_YLA_TE_X

Language Science Press

xHain

Grünberger Str. 16

10243 Berlin, Germany

<http://langsci-press.org>

Storage and cataloguing done by FU Berlin

Freie Universität  Berlin

Contents

1	Introduction	
	Angela Bartens, Peter Slomanson & Kristoffer Friis Bøegh	1
I	Creolized languages in the circum-Caribbean region	
2	African ethnolinguistic diversity in the colonial Caribbean: The case of the eighteenth-century Danish West Indies	
	Kristoffer Friis Bøegh	13
3	The expression of motion events in Haitian Creole	
	Carolin Ulmer	65
4	Postulating Atlantic English Pidgin/Creole as a pluriareal language: A perception study	
	Angela Bartens, Kwaku Owusu Afriyie Osei-Tutu & Tamirand Nnena De Lisser	97
II	Partially restructured varieties in the circum-Caribbean region	
5	A Gbe substrate model for discontinuous negation in Spanish varieties of Chocó, Colombia: Linguistic and historical evidence	
	Eliot Raynor	133
6	The minutes of the <i>Irmandade Nossa Senhora do Rosário dos Homens Pretos</i> in Recife: A case study	
	Fernanda M. Ziober	157
7	Language contact in Puerto Rico: Documenting an emerging variety of English	
	Sally J. Delgado	177
	Index	212

Chapter 1

Introduction

Angela Bartens^{a,b}, Peter Slomanson^{c,a} & Kristoffer Friis Bøegh^{d,e}

^aUniversity of Turku ^bUniversity of Helsinki ^cTampere University ^dUtrecht University ^eAarhus University

New research on circum-Caribbean creoles and language contact was first intended as a collection of sociolinguistically and sociohistorically oriented papers presented at the 2017 Summer Meeting of the Society for Pidgin and Creole Linguistics (SPCL), hosted by the Universities of Tampere and Turku. The contributors have since expanded the content of their original papers substantially, which has contributed to the empirical and analytical depth of their submissions.

The volume is organized as follows. The first part features three papers on language contact and creolized languages in the traditional, albeit now frequently debated sense (cf. Holm 1989, McWhorter 1998, DeGraff 2020). While we acknowledge this ongoing debate (cf. also McWhorter 2018 for a monograph treatment of the topic), we do not aim to engage with it directly in the present volume. The second part features three papers on partially restructured varieties (cf. Holm 2004) of the circum-Caribbean region (Fleischmann 1986).

By comparing the sociohistorical and current sociolinguistic outcomes of language contact in the circum-Caribbean region, the present volume introduces new comparative research and constitutes an invitation to further inquiry, both historical and contemporary, along similar lines. What sets of circumstances, past and present, can be claimed to obtain in contact situations, and how do those circumstances contribute to linguistic outcomes and to the ways in which contact varieties are viewed by their speakers?

Part I of the volume starts with KRISTOFFER FRIIS BØEGH's paper "African ethnolinguistic diversity in the colonial Caribbean: The case of the eighteenth-century Danish West Indies". In this in-depth and well-documented study, Bøegh



reconstructs a profile of the African-origin ethnolinguistic groups present in the Danish West Indies (now the US Virgin Islands) during the eighteenth century, based on the triangulation of four bodies of direct and indirect evidence of African languages in the colony. His findings have implications not only for creolistics, but contact linguistics in general, as they shed new light on the sociohistory of high-contact settings, such as the circum-Caribbean region.

The Danish West Indies constitute an interesting case for the study of African origins and ethnolinguistic diversity in the Caribbean. Overall, there is little direct linguistic evidence in the form of quotes or texts in African languages from the Caribbean during the colonial period. In the case of the Danish West Indies, however, documentation on both the transatlantic and intra-Caribbean slave trade enables the author to trace a detailed picture in the sense of Mufwene's (1996) Founder Principle. Whereas the Trans-Atlantic Slave Trade Database and its different versions have, at times, been criticized (cf. Inikori 2011), it constitutes an apt tool for reconstructing transatlantic and intra-Caribbean forced population movements.

In addition, Bøegh is able to draw on materials compiled by the Moravian missionaries, who were also active in other regions, such as the Nicaraguan Miskito Coast (cf. the establishment of the Herrnhut/Moravian Mission in Bluefields in 1848). These materials include Seidel's 1753 list of Christianized Africans and African Caribbeans from St. Thomas, baptismal records from the Friedenthal mission station on St. Croix (covering the period 1744–1832), and Oldendorp's linguistic work on both of these islands (Oldendorp 1777, 2000, Pope 1970, Seidel 1753).

The deep concern of the Moravian missionaries with language (cf. Holm 1989: 17–20) has provided us with significant data from the creole-speaking (and other linguistic) regions, best recognized in the field of missionary linguistics where, however, early language descriptions tend to look at the languages in question through the Greco-Latin grammar tradition.

Summarizing, the quantitative and qualitative data collected by Bøegh point to greater ethnolinguistic variety than was previously assumed to have existed in the Danish West Indies. This scenario is potentially applicable to other Caribbean and high-contact regions. The study definitely highlights the significance of the maintenance of both historical and present-day multilingualism as well as *lingua francas* in the New World context (cf. Álvarez López 2004, Palmié 2006).

Talmy's typology (1991) distinguished between the ways that motion events are framed in different languages, with Germanic languages as satellite-framed and Romance languages as verb-framed in their expression of the path of motion (i.e., the English "Mary danced into the room", as opposed to the dispre-

ferred “Mary entered the room dancing” in which the path is conveyed by the verb itself). In her paper in this volume, “The expression of motion events in Haitian Creole”, CAROLIN ULMER investigates how motion events are expressed by a group of native Haitian Creole speakers living in Germany. Since Kwa languages and French were component languages in the genesis of Haitian, the fact that the Kwa languages use serial verbs to express not just path, but motion as well, raises the question of what these competing diachronic influences have yielded in the grammar of Haitian and in the Haitian of the group of expatriate speakers that Ulmer worked with. This is an interesting and important question, not just for Haitian, but for the Atlantic creoles generally, including the circum-Caribbean group, given substrate influences from African languages with serial verbs and superstrate influence from European languages that are not similarly equipollently framed, the descriptive term employed by Slobin (2004), to characterize this third language type.

In “Postulating Atlantic English Pidgin/Creole as a pluriareal language: A perception study”, ANGELA BARTENS, KWAKU OWUSU AFRIYIE OSEI-TUTU, and TAMIRAND NNENA DE LISSER explore the idea of mutual intelligibility involving Atlantic English Pidgin/Creole (EP/C) from the point of view of a pluriareal, as opposed to a pluricentric language. (Note that Atlantic EP/C is distinguished from Atlantic EP/Cs. The authors use the latter to refer to distinct English-lexifier languages of the Atlantic region.) Despite the fact that some scholars, e.g., Dollinger (2019), find the distinction unnecessary, the authors argue for a hypothetical step from a pluriareal to a pluricentric language.

To merit the label ‘pluricentric’ a language has to comply with certain criteria, e.g., 1) being used in at least two nations or “interacting centers”; 2) having official status at state or regional level; 3) manifesting linguistic distance in the sense of Kloss’s (1967) *Abstand* languages as a symbol of linguistic and cultural identity; 4) enjoying acceptance as a pluricentric, but also national/regional variety; 5) being codified or in the relevant process and thence disseminated, e.g., through formal education (cf. Muhr 2012, Clyne 1992).

The term “pluriareal languages”, on the other hand, focuses on linguistic differences in these language varieties independent of national and political borders (Niehaus 2015, Elspaß et al. 2017). The authors argue that Atlantic EP/C might be on its way from a pluriareal to a pluricentric language, a point of view basically shared by Faraclas (2020).

In the first part of this study, 56 Ghanaian, Guyanese, and Nigerian informants were presented with audio samples of Ghanaian, Nigerian, Jamaican, and Sranan EP/C. The second part consisted of 20 interviews of Atlantic EP/C speakers con-

ducted in Guyana. Besides researching mutual intelligibility, the authors were interested in the acceptability of a common writing system.

Part II of the volume begins with ELIOT RAYNOR's paper "A Gbe substrate model for discontinuous negation in Spanish varieties of Chocó, Colombia: Linguistic and historical evidence". This study revisits so-called "double" or discontinuous negation (henceforward NEG2) and the role of African substrate languages, in this case in Chocó Spanish. Chocó is a Colombian department on the northwestern Pacific Coast of Colombia mainly inhabited by Afro-descendants (82.1%). Chocó Spanish differs from other Colombian varieties, e.g., in employing the linguistic structure studied in this contribution. In NEG2, an utterance-final negative element occurs in the proposition in addition to the preverbal negator. It is important to note, however, that the use of this construction is not canonical and actually stigmatized, albeit very commonly used, according to Raynor's informants. Despite the fact that no balanced sociolinguistic corpus of Chocó Spanish exists, Ruíz-García (2001) showed in a previous study that the same pragmatic conditioning found in Brazilian Portuguese, i.e., NEG2 in occurrences of the negation of a proposition activated in prior discourse, does not hold for Chocó Spanish.

Frequently, substratist approaches to Atlantic Creole and African-influenced varieties of the Americas ignore local adstrate languages. Raynor nevertheless shows that, at least in the case of NEG2, the indigenous languages of the Chocó are not candidates for its occurrence in Chocó Spanish. The terrain for the discussion of African influence is prepared through a meticulous presentation of the slave groups present in the region during the late 17th and early 18th centuries, the period postulated for the emergence of NEG2 in Chocó Spanish. Based on these data, Raynor considers that Gbe speakers, especially Ewe and Gen, were the key agents in its adoption as *bozales*, i.e., African-born slaves, outnumbered locally born *criollo* slaves up to at least 1778. The robustness of his findings on the African origin of the construction versus an Ibero-Romance one (cf. Sessarego 2017) is important for the discussion of the same in other Afro-American contact varieties. In addition, recalling that the ethnonym *Mina* should not be categorically associated with Elmina Castle in Ghana opens new perspectives in creolistics.

In "The minutes of the *Irmandade Nossa Senhora do Rosário dos Homens Pretos* in Recife: A case study", FERNANDA M. ZIOBER examines 32 nineteenth-century texts, transcribed by herself following paleographic methodology. 13 of the texts were written by Manoel de Barros, a member of the Brotherhood of Our Lady of the Rosary of Black Men, whereas the remaining texts feature three different calligraphies. Written by persons with limited access to formal education, the

texts show interesting deviations from the Portuguese orthography of the time. These deviations reveal, for example, phonetic and morphosyntactic phenomena, among them metathesis, shared by other Portuguese varieties, e.g., Angolan and Mozambican Portuguese (Petter 2009), cf. *pratido* < *partido* ‘broken’, and variable agreement within the NP, cf. *dos homem pretos* < *dos homens pretos* ‘two black men’.

As the author points out, this type of text constitutes an important source for the historical comparison of writing and, through that lens, Portuguese language as practiced by less literate people both in Brazil and in Africa during the nineteenth century, so crucial for, at least, the formation of Brazilian Portuguese (see Noll 1999).

In the third contribution of Part II, “Language contact in Puerto Rico: Documenting an emerging variety of English”, SALLY J. DELGADO investigates Puerto Rican English as an emerging contact variety spoken in Puerto Rico, one with which its speakers have come to identify, in appreciation of its local character. The author addresses the controversy surrounding the term “Spanglish” in previous literature as a cover term for language mixing, a term that has sometimes been used to denigrate the linguistic behavior of Spanish-English bilinguals in the Caribbean and beyond. She introduces the component activities of the piloted data collection project, involving the collection of spoken language, written language, and the documentation of language attitudes. She proposes that involving speakers in the creation of data resources for this variety would serve to increase awareness of its value, its status, and its function. This goal goes beyond the use of native speakers to guarantee the authenticity of the data for the sake of scientific work, but aiming to accomplish an important social goal for the long-term benefit of the speech community.

The awareness that this work and its dissemination can produce is critical to the emotional health of people who should be benefiting from the rich and well-defined array of linguistic resources available to speakers in bilingual cultural communities such as Puerto Rico. This is significant because making use of those resources in such a cultural context, while keeping the associated codes (i.e., Spanish and English) discrete, would be socially and communicatively dysfunctional. Unfortunately, the culturally appropriate non-discrete use of these resources is often impeded by internalized stigma surrounding language mixing and the effects of language contact generally. Delgado concludes that speakers would benefit from the increased presence of translanguaging in formal educational contexts, rather than restricting language mixing to informal contexts in which it is conducive to deeper interpersonal communication, social bonds, and the resulting collective emotional health of the speakers.

The subfield of creolistics has become broader and more inclusive in recent years, so that contact languages that cannot be characterized as pidgins and creoles are of increasing interest to creolists. We are very much in favor of this development. At the same time, we retain an interest in the specific circumstances and linguistic ecologies found in the Caribbean and sociohistorically-related areas in other parts of the Americas, from the period in which these areas were viewed in a sense as core target areas for research on creolization (and decreolization), based on the features of their contact languages, as well as on perceived cross-linguistic parallels between those languages. Those parallels included features such as similar periphrastic tense, mood, and aspect marking systems, verb serialization, syllable structure, and others that were observed across lexical inventories (i.e., English, French, Spanish, Portuguese, and Dutch). This interest is not intended as a prioritization of the circum-Caribbean region above other research areas within contact language research, but rather as a validation of the new work that has continued to appear on the fascinating language varieties that developed in the region and that continue to thrive there.

Acknowledgements

New research on circum-Caribbean creoles and language contact is a project that started with papers presented at the 2017 Summer Meeting of the Society for Pidgin and Creole Linguistics (SPCL), hosted by the Universities of Tampere and Turku. Its progress was delayed at times by obstacles such as the Covid-19 pandemic and the reorganization of teaching, research, and everyday life that resulted from it. Additionally, a number of technical roadblocks arose that delayed completion of the project.

Consequently, we feel enormous gratitude towards the authors of the papers in this volume. Throughout the process, their faith in it was unwavering and their patience went even further.

This volume could not have been completed without the generous assistance of twelve anonymous reviewers from a geographic area that extends from Haiti, Brazil, the United States, and Ghana to Finland, as well as the mentors the authors have cited in their individual acknowledgments.

We would also like to thank the community proof readers for their input, Maria Pesola-Gallone for some initial formatting, and Felix Kopecky for final typesetting. The Society for Caribbean Linguistics accepted to partially fund the costs associated with the typesetting which we hereby graciously acknowledge.

Furthermore, we wish to express our gratitude to Language Science Press in general, and first and foremost to Joseph T. Farquharson, Chief Editor of the Studies in Caribbean Languages series, for accepting this volume for publication.

Most importantly, however, we would like to underline the pivotal role played by Kristoffer Friis Bøegh. The volume was conceptualized and primarily edited by Angela Bartens and Peter Slomanson. Kristoffer Friis Bøegh joined as the third editor quite late in the process, substantially contributing to its finalization and ensuring a timely publication of the long-awaited volume.

References

- Álvarez López, Laura. 2004. *A língua de camões com iemanjá: Forma e funções da linguagem do candomblé*. Department of Spanish, Portuguese & Latin American Studies, University of Stockholm. (Doctoral dissertation).
- Clyne, Michael. 1992. Pluricentric languages: Introduction. In Michael Clyne (ed.), *Pluricentric languages: Differing norms in different nations*, 1–10. Berlin: De Gruyter.
- DeGraff, Michel. 2020. Toward racial justice in linguistics: The case of Creole studies. *Language* 96(4). 292–306. DOI: 10.1353/lan.2020.0080.
- Dollinger, Stefan. 2019. *The pluricentricity debate: On Austrian German and other Germanic standard varieties*. Milton Park: Routledge Focus.
- Elspaß, Stephan, Christa Dürscheid & Arne Ziegler. 2017. Zur grammatischen Pluriarealität der deutschen Gebrauchsstandards – oder: Über die Grenzen des Plurizentritätsbegriffs. *Zeitschrift für Deutsche Philologie* 136. 69–91.
- Faraclas, Nicholas. 2020. Identity politics. In Miriam Meyerhoff & Umberto Ansaldo (eds.), *Routledge handbook on Pidgin and Creole languages*, 269–285. Abingdon: Routledge.
- Fleischmann, Ulrich. 1986. *Das Französisch-Kreolisch in der Karibik: Zur Funktion von Sprache im sozialen und geographischem Raum*. Tübingen: Gunter Narr.
- Holm, John. 1989. *Pidgins and Creoles*, vol. 2. Cambridge: Cambridge University Press.
- Holm, John. 2004. *Languages in contact: The partial restructuring of vernaculars*. Cambridge: Cambridge University Press.
- Inikori, Joseph E. 2011. *Extending the frontiers: Essays on the new Trans-Atlantic Slave Trade Database*. David Eltis & David Richardson (eds.). New Haven, CT & London: Yale University Press. 377.
- Kloss, Heinz. 1967. “Abstand languages” and “Ausbau languages”. *Anthropological Linguistics* 9(7). 29–41.

- McWhorter, John H. 1998. Identifying the Creole prototype: Vindicating a typological class. *Language* 74(4). 788–818. DOI: 10.2307/417003.
- McWhorter, John H. 2018. *The creole debate*. Cambridge: Cambridge University Press. DOI: 10.1017/9781108553308.
- Mufwene, Salikoko S. 1996. The founder principle in creole genesis. *Diachronica* 13(1). 83–134. DOI: 10.1075/dia.13.1.05muf.
- Muhr, Rudolf. 2012. *Non-dominant varieties of pluricentric languages getting the picture*. In collaboration with Catrin Norrby, Leo Kretzenbacher, Carla Amorós Negre. In memory of Michael Clyne. Frankfurt am Main: Peter Lang.
- Niehaus, Konstatin. 2015. Areale Variation in der Syntax des Standarddeutschen: Ergebnisse zum Sprachgebrauch und zur Frage Plurizentrik vs. Pluriarealität. *Zeitschrift für Dialektologie und Linguistik* 82(2). 133–168.
- Noll, Volker. 1999. *Das brasilianische portugiesisch: Herausbildung und Kontraste*. Heidelberg: Winter.
- Oldendorp, Christian Georg Andreas. 1777. *Geschichte der Mission der evangelischen Brüder auf den caraibischen Inseln S. Thomas, S. Croix und S. Jan*. J. J. Bossart (ed.). Barby: Laux.
- Oldendorp, Christian Georg Andreas. 2000. *Historie der caribischen Inseln Sanct Thomas, Sanct Crux und Sanct Jan, insbesondere der dasigen Neger und der Mission der evangelischen Brüder under derselben. Erster Teil*. Hartmut Beck, Gudrun Meier, Stephan Palmié, Peter Stein & Horst Ulbricht (eds.) (Abhandlungen und Berichte des Staatlichen Museums für Völkerkunde Dresden 51). Berlin: VWB Verlag für Wissenschaft und Bildung.
- Palmié, Stephan. 2006. Creolization and its discontents. *Annual Review of Anthropology* 35(1). 433–456. DOI: 10.1146/annurev.anthro.35.081705.123149.
- Petter, Margarida. 2009. O continuum afro-brasileiro do português. In Charlotte Galves, Helder Garmes & Fernando Rosa Ribeiro (eds.), *África-Brasil: Caminhos da língua portuguesa*, 159–174. Campinas: Editora da Unicamp.
- Pope, Pauline Holman. 1970. *Cruzan slavery: An ethnohistorical study of differential responses to slavery in the Danish West Indies*. Davis: University of California, Davis. (Doctoral dissertation).
- Ruíz-García, Martha. 2001. *El español popular del chocó: Evidencia de una reestructuración parcial*. Albuquerque, NM: University of New Mexico. (Dissertation).
- Seidel, Nathanael. 1753. *Zeugniße von den Communicanten unser Neger-Gemeine in S. Thomas verfertiget von Bruder Nathanael in Gemeinschaft der dortigen Arbeiter bey seiner Visitation in S. Thomas im Monat Junio und Julio 1753*. Manuscript.

- Sessarego, Sandro. 2017. Chocó Spanish double negation and the genesis of the afro-Hispanic dialects of the Americas. *Diachronica* 34(2). 219–252. DOI: 10.1075/dia.34.2.03ses.
- Slobin, Dan. 2004. The many ways to search for a frog: Linguistic typology & the expression of motion events. In Sven Strömquist & Ludo Vorhoeve (eds.), *Relating events in narrative: Typological and contextual perspectives*, vol. 2, 219–257. Mahwah, NJ: Lawrence Erlbaum Associates.
- Talmy, Leonard. 1991. Path to realization: A typology of event conflation. In Laurel Sutton, Christopher Johnson & Ruth Shields (eds.), *Proceedings of the seventeenth annual meeting of the Berkeley Linguistics Society*, 480–519. Berkeley: Berkeley Linguistics Society.

Part I

Creolized languages in the circum-Caribbean region

Chapter 2

African ethnolinguistic diversity in the colonial Caribbean: The case of the eighteenth-century Danish West Indies

Kristoffer Friis Bøegh

Utrecht University & Aarhus University

This study reconstructs a profile of the ethnolinguistic groups originating in Africa represented in the eighteenth-century Danish West Indies (i.e., St. Thomas, St. John, and St. Croix, today's US Virgin Islands), based on triangulation of four bodies of direct and indirect evidence of African languages in the colony. I analyze i) data on the recorded slave trade to the colony; ii) data on 26 African languages collected in peer-group interviews conducted in the late 1760s by the German Moravian missionary and historian C.G.A. Oldendorp, with an estimated 70 enslaved individuals on St. Thomas and St. Croix; iii) data on the “nations” (i.e., approximate ethnic backgrounds) represented in a sample of 418 Christianized Africans and African Caribbeans of the Moravian congregation on St. Thomas, recorded in 1753 by the missionary Nathanael Seidel; and iv) baptismal records from 1744–1832 collected by the Moravians on St. Croix, featuring systematic information on more than 6,000 baptismal candidates of African descent. The results show that a wide range of languages were spoken within the enslaved population. In addition, a number of differences are identified in the distribution of groups and their proportions between St. Thomas and St. Croix. The findings, I argue, have implications by extension for contact linguistics more generally, shedding new light on the sociohistorical settings characterizing high-contact sites in the circum-Caribbean region.

1 Introduction

During the seventeenth and eighteenth centuries, European colonial powers divided the Virgin Islands in the Caribbean into two primary geopolitical units:



the British Virgin Islands and the Danish West Indies.¹ The latter group, which became part of the United States as the US Virgin Islands in 1917, comprises St. Thomas (permanently settled in 1672), St. John (1718), and St. Croix (1734). As in the case of many other colonies in the circum-Caribbean region, the developmental history of the Danish West Indies was shaped by trade, plantation slavery, and contacts between diverse population groups of predominantly African and European origins (for an overview, see, e.g., Jensen & Simonsen 2016). Although the islands were Danish by name for centuries, the vast majority of their population did not originate from Denmark² but had African or African Caribbean backgrounds. The enslaved Africans transported to the islands carried with them a large measure of their cultural systems, including their languages; and because they originated from numerous localities, an ethnolinguistically diverse society formed early on, and continued throughout the period of slavery. The colonizing population was linguistically heterogeneous as well. However, whereas information on the different backgrounds represented in the latter group is readily accessible (including in the form of copious census data), some of the central primary sources on the provenance and linguistic backgrounds of the enslaved population have yet to be brought together in an investigation of the ethnolinguistic diversity of people of African descent represented in the slave society.

This study aims to reconstruct a profile of the ethnolinguistic groups originating in Africa represented in the eighteenth-century Danish West Indies. In addition, it seeks to uncover differences in the distribution of groups and their proportions between the colony's two main islands, St. Thomas and St. Croix, which were settled around half a century apart, and whose demographic profiles, consequently, reflect both differences in local developments and shifting waves in the slave trade. Moreover, because of the general scarcity of documentary evidence of African languages in the context of the colonial-period Caribbean, the study also has a wider relevance, extending beyond the scope of Virgin Islands language history. There is a seemingly unparalleled volume and breadth of documentation from the Danish West Indies relevant to a case study nuancing current notions about the extensiveness of diversity in the ethnolinguistic composition of "typical"³ Caribbean plantation slave societies. From a contact linguistics perspective, tapping into this documentation has the potential to cast new light on

¹In addition, the islands of Vieques and Culebra, located just east of Puerto Rico and both designated as municipalities of that territory, are also geographically part of the Virgin Islands.

²From 1672 to 1814, the Danish West Indies belonged to the double monarchy of Denmark-Norway, after which time Norway gained independence from Denmark. In the present study, I refer to the double monarchy simply as "Denmark."

³In a classic study of the Danish West Indies under Company rule, Westergaard (1917: 121) characterized the colony as a "fairly typical plantation society."

2 African ethnolinguistic diversity in the colonial Caribbean

the sociohistorical settings and hence contributions of agents involved in the development of creole languages in the region – provided the situation characterizing the Danish West Indies is comparable to those characterizing other territories. This is most likely the case; although slavery and settlement varied considerably across the Caribbean, there are also many parallels from territory to territory (see, e.g., Higman 2011). Against this backdrop, the present study can be viewed as a response to calls from within the field for studying language contact phenomena, above all else, from a historically realistic perspective (along the lines of, e.g., Arends 2008; cf. also Kouwenberg & Singler 2018).

The Danish West Indies occupy a special position in the context of documentation of African languages in colonial settings. In addition to data on the recorded slave trade (the existence of which is commonplace throughout the Caribbean), there is an abundance of direct and indirect evidence of African languages, among others, in the form of data recorded *in situ* by the Moravian Brethren, the pietistic missionary movement active in the Danish West Indies since 1732 (cf. §2.3 for details). Focusing on the eighteenth century (and for some aspects extending into the first part of the nineteenth century), I base my analysis primarily on triangulation of the following four bodies of evidence:

- data on the recorded transatlantic and intra-Caribbean slave trade to the Danish West Indies (including records from the online repository available at <https://www.slavevoyages.org/>, henceforth the Voyages Database);
- direct and indirect evidence of 26 African languages collected in peer group interviews with 70 or so enslaved individuals on St. Thomas and St. Croix, conducted by the Moravian missionary C.G.A. Oldendorp in the late 1760s for an ethnographic survey (Oldendorp 1777; English edition: Highfield & Barac 1987; critically-edited unabridged version: Oldendorp 2000, 2002);
- self-reported data on the “nations” (i.e., approximately, ethnic backgrounds) from a sample of 418 Christianized Africans and African Caribbeans of the Moravian congregation on St. Thomas, recorded in 1753 by the missionary Nathanael Seidel (Seidel 1753; list published as an appendix in Sebro 2010);
- baptismal records for the period 1744–1832 from Friedensthal, the oldest of the Moravian mission stations in St. Croix, featuring systematic information on the ethnic backgrounds of over 6,000 baptismal candidates with an African or African Caribbean background (tabulated in Pope 1970).

Additionally, where relevant, I draw on the historical literature on the Danish West Indies, including Danish-language research based on sources from the Danish National Archives (cf. Jensen & Simonsen 2016 for an overview).

The study is organized as follows. In §2, I offer background information on the Danish West Indies and discuss the colony's suitability as a case study on African ethnolinguistic diversity in the colonial Caribbean setting. Moreover, I contextualize the sources of evidence I use to trace members of the African population to specific areas and ethnolinguistic groups. In §3, I examine the recorded slave trade to the Danish West Indies, offering an assessment based on estimates of imports from both the transatlantic and intra-Caribbean slave trade. Based on this exercise, I argue for a need to consider additional sources of evidence so as to be able to qualify the big-picture result which obtains. In §4, accordingly, I shift perspective to locally-recorded evidence. I compare linguistic data (as well as other cultural evidence) collected by Oldendorp with modern, carefully transcribed African language data, aiming to identify present-day correlates to the languages documented in the historical material. Extrapolating from the results of the analysis of Oldendorp's data, I then examine Seidel's list of Christianized Africans and African Caribbeans from St. Thomas in §5 and the baptismal records from St. Croix in §6, looking at similarities and differences in the distribution of the identified ethnolinguistic groups and their proportions between the two islands. In §7, finally, I provide a summary of the ground covered in the study and present some concluding remarks.

2 The Danish West Indies as a case study on African origins and ethnolinguistic diversity: Background and sources of evidence

Setting the scene, this section presents background information on the Danish West Indies and its enslaved population, focusing on the setting in the eighteenth century. I discuss the problem of ethnolinguistic identifications in the colonial Caribbean, and I contextualize the various sources of evidence drawn upon in the study.

2.1 The slave society in the eighteenth century: Demography and ethnolinguistic diversity

In 1672, the Danish West India and Guinea Company founded a permanent colony on the island of St. Thomas. The aim was to follow the example set by other

2 *African ethnolinguistic diversity in the colonial Caribbean*

expansionist European nations that had established exploitation colonies overseas in order to gain a part of the wealth associated with the transatlantic trade (cf. Jensen & Simonsen 2016; see also Hall 1992; Gøbel 2016; Olsen 2017; Roberts et al. 2024, among other works). St. Thomas is hilly and was not particularly well suited for large-scale plantation agriculture, but it is endowed with an excellent natural harbor and, therefore, gained importance as a transshipment port over the course of the eighteenth century. In 1718, the Danish colony expanded with the annexation of the unsettled neighboring island of St. John. In 1733–34, more significantly, St. Croix was acquired from the French, and its colonization was initiated. The topography of St. Croix was well suited to plantation agriculture, and as a result, following its incorporation into the colony, sugar cultivation increasingly became an economic mainstay of the Danish West Indies. This relied on highly labor-intensive cultivation techniques and involved an ever-growing demand for an imported workforce: the enslaved Africans and African Caribbeans. The following section will consider the historical demography of the enslaved population across the three islands.

2.1.1 **Slave demography in St. Thomas, St. John, and St. Croix**

In its first century, the Danish colony was characterized by a rapidly growing enslaved population. As early as 1686, fourteen years after the permanent settlement of St. Thomas, people of African descent exceeded the European population, and that development only continued to accelerate throughout the eighteenth century. Thus, in 1755, when the Danish Crown took over the colony from the Company, there were 1,760 inhabitants of European origin across the three islands combined, whereas the enslaved numbered 14,409 people, or nearly 90% of the total population (Gøbel & Sebro 2017: 58). However, the demographic development was not the same on the three islands.

The main population growth on St. Thomas occurred in the first two decades of the eighteenth century. After that, the number of enslaved people remained stable for the remainder of the century. Based on an analysis of the colony's extant tax-rolls, Green-Pedersen (1971: 157) reported the following population figures. In 1686, the colonizing population numbered 300 and the colonized 333; by 1720, the numbers were 524 against 4,504; by 1754, they were 228 against 3,481 (with this decline in absolute numbers being offset by an increase on St. Croix). The European population segment comprised 7% of the island's total number of people in 1755, just under 8% in 1775, and some 15% in 1803.

St. John was settled from St. Thomas, and it developed largely as an adjunct of that island. Its colonization proceeded rapidly until 1733, at which time it became

the scene of a large-scale slave insurrection. The rebellion was put down only after six months, and the island remained sparsely inhabited thereafter (Hornby 1980: 91–96). There were 2,031 enslaved people on St. John in 1755, corresponding to approximately 90% of the island’s population. This number fluctuated a bit in the latter eighteenth century, rising to 2,598 in 1803, after which it remained stable (Green-Pedersen 1971: 157–58; Gøbel & Sebros 2017: 59).

St. Thomas’ economy relied much less on plantation agriculture than those of St. John and, in particular, St. Croix. This is reflected in the fact that, peaking in 1797, as many as 34% of the enslaved on St. Thomas worked in the town of Charlotte Amalie (Green-Pedersen 1981: 234). Sebros (2016: 128) notes that society in St. Thomas “was based on the principally racial divide where black meant slave and white meant free.” However, as in slave societies elsewhere, this principal divide was not absolute. From the very beginning, a small group of free African Caribbeans was present as well. They managed in different ways to find space for themselves in the slave society, especially in the urban setting, by negotiating ideas of race and social position. However, in the eighteenth century, St. Thomas was still in the early stages of that process. Therefore, the free African Caribbean population, Sebros (*ibid.*) indicates, “was structurally of no great importance.” The same was the case on early Danish St. Croix. Urban slavery also became firmly established there early on (see Tyson 2011), but by far the largest proportion of the enslaved on St. Croix spent their days on large-scale rural sugar plantations. In 1792, for example, 86% of enslaved people on St. Croix were engaged in field labor (Hall 1992: 75).

St. Croix was surveyed and brought under European cultivation in the course of the 1730s and 1740s, and the island’s plantation sector was consolidated in the 1750s (Hopkins 1987). The following years were characterized by a rapid growth in the enslaved population, and by high mortality rates. Approximately 1% more died than were born (Simonsen & Olsen 2017: 150), meaning that a constant inflow of new arrivals was needed to maintain existing numbers, not to mention meeting the requirement for absolute growth (cf. Roberts et al. 2024). In terms of demography, St. Croix thus followed the developmental pattern typical of a Caribbean plantation colony (cf. Higman 2011: 130). According to Simonsen & Olsen (2017: 150–151), in 1736, two years after the island’s settlement, there were 137 enslaved people on St. Croix; in 1742, this number had risen to around 1,700; and in 1765, it had increased to almost 19,000. This growth continued into the 1770s, after which the number stabilized at around 23,000 or 24,000. Between 1793 and 1803, the enslaved population grew once again, peaking at approximately 27,000 people in 1804. After that, the number decreased steadily. The enslaved claimed their freedom in 1848. In 1847, the last year a census was taken of the

enslaved population, there remained 16,673 people (Holsoe 1994: 33), or about 62% of the all-time maximum recorded 44 years earlier.

2.1.2 Ethnolinguistic diversity in the Danish West Indies: A tentative orientation

From the outset, Danish colonization in the Caribbean was characterized by a lack of much of an actual Danish colonizing population. This resulted in what has been described as a *de facto* policy of “colonization by invitation” (Hall 1992: 6), which gave rise to substantial numbers of foreign Europeans being introduced into the population. These people were particularly of Dutch heritage on St. Thomas and St. John (Arends & Muysken 1992: 51), and of British stock on St. Croix (Hall 1992: 13). As for the languages used among the Europeans, Oldendorp (2000: 357) comments, writing in the 1760s, that at least English, Dutch, Dutch Creole (more on this below), German, Danish, French, and Spanish were in use. Although considerable, this diversity did not match that represented by the African languages that were in use in the colony. By way of illustration of this point, Johann Gottlieb Miecke, the principal of the Moravian Mission in St. Croix in the late eighteenth century, wrote in his 1796 letter of resignation (quoted in Lawaetz 1902: 124) about the linguistic diversity characterizing the island. Miecke stated that he was losing faith in the mission work since it would be insufficient, he felt, to know even as many as 20 different languages if one was to effectively reach the enslaved population. Thus, the eighteenth-century context of the Danish West Indies was clearly characterized by extensive multilingualism, both among the enslaved and colonizing subsets of the population.

Meanwhile, this sociohistorical setting had led to the establishment of creole languages with lexical bases in Dutch and English, respectively, in the colony.⁴ The first of these, Virgin Islands Dutch Creole, most likely formed locally on St. Thomas around 1700 (cf. Sabino 2012: 52–54). It subsequently gained currency as a lingua franca in the colony, especially on St. Thomas and St. John. This language was the first creole to be described in a printed grammar (see Magens 1770), and it is also notable for having been documented extensively for more than 250 years (for an anthology of Virgin Islands Dutch Creole texts, see van Rossem & van der Voort 1996; see also van Rossem 2017; van Sluijs 2017; Bøegh et al. 2022, along with other recent works). By the second half of the 1700s, a

⁴In addition, Highfield (1993: 131–32) conjectured that an undocumented French-lexifier creole was in use in St. Croix prior to the island’s abandonment by the French in 1696. Whether or not such a variety did exist, it is not to be confused with the French dialect of Frenchtown, St. Thomas (described in Highfield 1979).

variety of Caribbean English Creole had become sufficiently well-established on St. Croix (where English had become established to the virtual exclusion of Dutch, cf. Oldendorp 2000: 358, 682) for contemporary observers to make reference to its existence (e.g., Auerbach 1774 quoted in van Rossem 2017: 90). English Creole gained ground on St. Thomas and St. John from the early 1800s onward, leading to the eventual decline of Virgin Islands Dutch Creole in the nineteenth century (as observed by, e.g., Pontoppidan 1881; cf. Stein 2014; for further discussion and documentary evidence, see, e.g., Bøegh & Bakker 2021).

The creoles underwent influence from the diverse African languages represented in the colony. As for the identities of the African languages, earlier studies supported the claim that most of the Africans brought to St. Thomas during the last part of the seventeenth century spoke Kwa languages such as Gã and Akan. Speakers of Akan, in particular, were estimated, by both historians and linguists, to have predominated throughout the peopling of St. Thomas (e.g., Feldbæk & Justesen 1980; Stolz & Stein 1986: 118; Parkvall 2000: 135). In addition, Sabino (1988) suggested that Ewe (Gbe subgroup of Kwa) was a main early substrate in the colony (cf. also Sabino 2012). More recent discussions of Danish-African trade relations and the transatlantic slave trade to the Danish West Indies (e.g., Sabino 2012: 64–67; van Sluijs 2017: 27–32) suggest that this picture is incomplete, in that a much larger set of languages were spoken in the regions from which enslaved people were transported to the Danish colony, and from a much wider area of Africa. Sabino (2012: 65) opts to focus on the Danish slaving operations as those that were most likely to have impacted the early demographic development of St. Thomas. Yet, the significance of other sources of trade to the growth of the colony's plantation sector should not be underestimated. For instance, in 1685, the North German Brandenburg Africa Company gained the right to use St. Thomas as a transshipment port, after which it supplied thousands of enslaved people to the colony, some of whom were not re-exported, but incorporated into the local slave society (Weindl 2008: 254–257), *pace* Sabino (2012: 65).

The most important previous analysis of slave imports with a focus on St. Croix is an unpublished anthropology dissertation (Pope 1970). That study's main finding was that the ethnic diversity on St. Croix was greater and somewhat differently composed than that of St. Thomas. The archival data undergirding the analysis in Pope (1970) are available for scrutiny in its appendices. In §6, I analyze these unique data anew, hence the short discussion here. Re-evaluation of the data is warranted seeing as how a number of Pope's (1970) conclusions have been subject to revision in later research (as detailed in §6).

In 1803, the official termination of the Danish slave trade effectively ended the flood of new contingents of native speakers of African languages. Asmussen

(1983: 48) calculated that 46.9% of the plantation laborers present on St. Croix in 1804 had been born in Africa, whereas only 7.3% had been in 1846. Thus, although the prospects for their maintenance and long-term survival were poor, we can assume that African languages were in (recessive) use well into the nineteenth century. Despite this, there has been only limited attention devoted in the literature to identifying the African languages. Besides the studies cited above, there are some previous tentative analyses of Oldendorp's material (Fodor 1975; notes from Peter Stein in Oldendorp 2000: 365–465; Jones 2010). A more precise overview of the ethnolinguistic diversity represented in the colony can be obtained through the use of new methods and materials.

2.2 The problem of ethnolinguistic identifications in the context of the colonial Caribbean

There is little direct linguistic evidence in the form of quotes or texts in African languages from the Caribbean in the colonial period. Outside of the Virgin Islands, there is some documentation from 1741 of a Gbe language in Brazil (da Costa Peixoto 1944), and from the late 1700s of Kikongo in Haiti (Peter Bakker, p.c.), but not much else has survived. The lack of direct evidence is a fact commonly acknowledged in creole studies, and one which routinely obstructs what Kouwenberg (2008: 2) has termed “[t]he first task of the substratist creole researcher” – namely, identifying as precisely as possible the input languages involved in a given contact setting so as to be able to apply historical-comparative methods to investigate the origins of the various creole languages as contact phenomena. On the other hand, there is abundant indirect evidence for the use of African languages in the Caribbean, including in the form of lexical Africanisms and other instances of substrate influence in today's creoles (see, e.g., Parkvall 2000, 2019; Bartens & Baker 2012; Muysken et al. 2015). Beyond this, however, evidence for the African input in sociohistorical linguistic research – such as in reconstructions of the settlement histories of individual creole-speaking societies – tends to be limited to metalinguistic observations coupled with documentation of slave voyages (see, e.g., Kouwenberg 2008, on Jamaica; Prescod & Fraser 2015, on St. Vincent; Ballester 2016, on Antigua, to mention but a few studies based primarily on such sources of evidence).

As discussed by Velupillai (2015: 109), archival information on the location of the trade forts and the points of embarkation for the slave ships has often been taken as an indication of the ethnolinguistic origins of substrate populations (cf. also §3). This has become possible because detailed information on the ports of embarkation, the number of captives shipped out, and their destinations has been

collected and made available for the transatlantic slave trade (see Eltis et al. 1999 and, especially, the updated online database based on this publication, available at <https://www.slavevoyages.org/>, i.e., the Voyages Database). On the downside, the location of the forts and the points of embarkation for the ships cannot always be taken to coincide with the origins of the enslaved individuals that those forts traded with. These may have been taken from different hinterland sources, or bought from traders from other regions. According to Velupillai (2015: 109), we should aim to establish the likely areas that would supply certain forts and ships, and which languages were likely to have been spoken in those areas at the time; yet, critically, as she notes (*ibid.*), “even if we were able to establish those facts, we would have extremely limited access, if any, to the kinds of varieties spoken in the given areas at the time.” In a good number of instances, however, this problem can be bypassed in the case of the Danish West Indies, as I will proceed to show in the present study.

Meanwhile, metalinguistic comments from contemporary observers – while also useful – likewise present us with a number of challenges. In the context of the Danish West Indies, this can be illustrated by the following quote from a police journal sourced from the Danish National Archives:

Carolina... Mandungo... She speaks no understandable language, and none of the black people in the [police] yard were able to understand her, so no one knew how she had come to the country. The police chief sent out the officers to summon various black people from the [West African] coast who were able to speak English, too [in addition to African languages, KFB]. After numerous attempts, a black boy, Peter, was summoned, who was able to understand the black woman. (Danish National Archives: Sheriff of Christiansted: Police journals, 3 September 1805, my translation)

Although this extract does present us with information relevant to the identification of the mentioned individual’s ethnic background (“Mandungo”), evidence of this kind is too limited to allow researchers to engage with questions pertaining to ethnolinguistic origins more generally. There is a lack of representativeness (in being anecdotal) and contextual information (on what basis did the record keeper reach the conclusion?), and it appears unclear whether there exists any direct evidence of the language used among the “Mandungo,” which, if present, could be compared with corresponding data from modern African languages to help establish its identity.⁵

⁵Oldendorp did in fact record “Mandungo” (or “Mandongo”) language data, which I track to either Guthrie’s (1948) Bantu Zone B or H, i.e., approximately between southern Gabon and northwestern Angola on the West-Central African coast (cf. §4.1.18).

2 African ethnolinguistic diversity in the colonial Caribbean

In light of the potential pitfalls of working with sources of evidence such as those discussed above, Prescod & Fraser (2015: 6), in their study on the settlement history of St. Vincent, sum up the main problem as follows: “The ethnolinguistic tapestry of the African continent is intricate and ascertaining ethnolinguistic origins is an arduous task.” Indeed, the African continent is characterized by vast linguistic diversity. The number of African languages is estimated to be over 2,000 in Eberhard et al. (2019), and the lexical diversity is enormous. In addition, typological unity cannot be attributed to the various families and subgroupings (e.g., Bøegh et al. 2016). Importantly, the availability of information on modern African languages for comparative research purposes has increased considerably over the course of the past few decades. African linguistic relationships are not yet fully delineated, though, and the internal structures and/or external affiliations of a number of groupings continue to be subject to regular proposals for revision (see, e.g., Güldemann 2018). As a point of departure, therefore, the present study adopts the broad classification of the *World Atlas of Language Structures* (Dryer & Haspelmath 2013) into linguistic genera, i.e., groups of languages “whose relatedness is fairly obvious without systematic comparative analysis” (Dryer 2005: 584). Where pertinent, more specialized divisions of groups commonly subsumed under the Niger-Congo phylum (or superfamily) follow the classification proposed by Williamson & Blench (2000: 18).

2.3 The Moravian witnesses to slavery in the Danish West Indies

In 1732, the Moravian Brethren – missionaries operating from Herrnhut, Germany – established their first overseas mission on St. Thomas, expanding to St. Croix in 1734 (see, e.g., Meier et al. 2010). They recorded the most central of the linguistic sources considered in the present study. This section details their activities in the Danish West Indies.

The Moravian presence in the colony was important in several respects. The missionaries were present from early in the colony, arriving 60 years after the settlement of St. Thomas, 14 years after the settlement of St. John, and at the very beginning of the Danish settlement of St. Croix. They were on a mission to Christianize the enslaved African labor force, and they were from the start looking to expand to other locations. The movement subsequently spread to Suriname (1735), Jamaica (1754), Antigua (1756), Barbados (1765), and elsewhere in the Caribbean and beyond. Thus, and seeing as how they were gathering experience for future expansion, the missionaries had embarked on a project that would periodically require evaluation. Accordingly, as noted by Highfield (1994: 141), they kept detailed records in the form of “[c]hurch books, journals, diaries, registers

of all kinds – but especially those relative to baptisms – correspondence, and reports,” all of which were carefully copied and preserved. The records were regularly inspected by visitants, whose task it was to supervise the progress made and to report back to Herrnhut. Their accounts, of which Oldendorp’s (1777/2000, 2002) is the best-known example, add important supplementary information to that collected on an ongoing basis in the colony, such as the baptismal records from St. Croix. All in all, the Moravians left an enormous amount of detailed information about their activities, and about the enslaved population in the Danish West Indies.

Initially, the Moravian effort was concentrated on St. Thomas. Later, that effort was largely redirected to St. Croix. After a difficult beginning in the 1730s and 1740s, the Moravians’ work in the islands succeeded remarkably in the second half of the eighteenth century.⁶ By 1768, they had baptized some 4,500 Africans and African Caribbeans; by 1794, that number had climbed to 9,300, and by 1815 it reached approximately 15,000 (Highfield 1994: 146–47). In 1804, by comparison, the total (and all-time maximum) number of enslaved people on St. Croix came to approximately 27,000 (Simonsen & Olsen 2017: 151). In each decade from 1760 to 1800, Simonsen (2017: 190) estimates, the Moravians baptized between 8% and 16% of the enslaved African population in St. Croix. By Highfield’s (1994: 147) estimate, the Brethren had, in effect, enrolled more than half of all the enslaved people in the Danish West Indies within approximately 70 years of the onset of their enterprise.

The success of the Moravians may be attributed largely to the fact that the enslaved had lost their kinship ties after having been captured, sold, and forcibly transported to the Caribbean. Sensbach (2005: 92–95) describes how the Moravians offered an alternative kinship system that appealed to the enslaved, and which afforded the members of the congregation certain opportunities and a chance for social elevation that was otherwise unavailable to those at the bottom of the slave society.

In the following sections, I will first consider quantitative evidence of the slave trade to the Danish West Indies, after which I turn to analyzing the various sources of documentation recorded by the Moravian Brethren. In this connection, it is important to note that I assume, though not uncritically, that Seidel’s list from St. Thomas, the baptismal records from St. Croix, and Oldendorp’s account based on slave interviews conducted on both islands are in general alignment with respect to their application of ethnic categories such as, e.g., “Kongo”

⁶During this period, a second, Danish-operated Lutheran mission was also established in the Danish West Indies (see, e.g., Larsen 1950). The Lutherans likewise left important documentation relevant to the study of Virgin Islands language history.

and “Amina,” and that the connection between such ethnic categories and their correlated languages is extrapolatable from the linguistic data presented by Oldendorp in his ethnographic survey. This assumption, corroborated by Simonsen (2017: 190–191), rests on the fact that the information considered was obtained directly from the people concerned. Oldendorp stated that his information was “based on their own narrations” (Highfield & Barac 1987: 159), often in the context of peer group interviews, on both St. Thomas and St. Croix.

3 Evidence of the transatlantic and intra-Caribbean slave trade to the Danish West Indies

3.1 The transatlantic slave trade and Danish-African trade relations

Following an unsuccessful experience with indentured labor (Hvid 2016), importation of enslaved people directly from Africa to the Danish West Indies began in the 1670s, most likely in 1673 (Westergaard 1917: 40). In total, the number of enslaved Africans exported by the Danes and through Danish establishments probably totaled no fewer than around 100,000 people (Gøbel 2016: 15). As the Danish slave trading forts were located on the so-called “Gold Coast” (present-day Ghana), and as the Danish West India and Guinea Company had been granted a national monopoly on shipping and trade to Africa and the West Indies (Gøbel 2016: 12), it would at first sight appear reasonable to assume that a preponderance of the enslaved were drawn from the ethnic groups of that region, where especially Akan lects are, and were, widespread (Hair 1967, 1968). From 1661, the Danish headquarters in Africa was Fort Christiansborg at Accra (Nørregaard 1966). Sabino (2012: 66) describes the Danes’ trade relations with local groups there. The Danes’ first trading partners there were the Gã, who were at war with the expanding Akan-speaking Akwamu. When slave import to St. Thomas commenced, the Gã were thus in a position to sell Akan-speaking prisoners of war from the interior. After defeating the Gã in 1681, the Akwamu began supplying the Danes with Gã prisoners, as well as with captives from western Gbe-speaking areas that they already dominated. The Danes continued to trade with the Akwamu until the 1730s. According to Parkvall (2000: 135), the situation sketched out here “eventually led to an exceptionally heavy Kwa bias in the workforce of the Danish West Indies.”

The reality was not as straightforward as that, however, in that even more enslaved people were drawn from other areas along the West African littoral, from Senegambia in the north to Angola in the south. Also, in the early period, many

of the captives landed in the colony were not necessarily from the Gold Coast region. Danish ships routinely took onboard captives from ports of call along the way to the forts, and these embarkation points were often not registered anywhere. Moreover, for a variety of reasons, the Company's purported monopoly on shipping and trade was not actually complied with, meaning that parties other than the Danes became involved in the trade (e.g., the earlier-mentioned Brandenburg Company). Thus, a Danish planter (quoted in Nielsen 1981: 83, my translation) wrote circa 1740 about the enslaved on St. Croix, that there were "as many nations [...] among them as there are landscapes, towns, and places in both America and especially Africa, where most of them hail from."⁷

The quote points to a mixed African-origin population, an observation that can be corroborated by considering data from the Voyages Database. A search for records of all ships with the Danish West Indies as the principal place of landing produced 381 records of voyages, representing a total of 89,829 disembarked individuals (105,253 embarked). Gøbel (2016: 15) notes that, despite allowing itself some extrapolation, the Voyages Database's total number of people embarked is in overall alignment with more specialized studies on the volume and composition of the slave trade to the Danish West Indies. Table 1 (page 28) shows an overview of extractable data on these voyages' embarkation regions. The figures are based on arrivals. The regions of origin are those of Eltis et al. (1999). Based on these data, a diversity of embarkation regions for the slave trade in the Danish West Indies can be observed. The figures are visualized in Figure 1.

Who were transported, from where they were transported, and by whom, was contingent on numerous factors, and it is evident from the slave trade data considered here that these factors varied over time. There would appear, indeed, to have been an early dominance of groups from the Gold Coast and the Bight of Benin – Akan, Gã, and Gbe speakers – and of people with unknown embarkation points, who, consequently, are subsumed under "Other Africa."⁸ West-Central Africa, where Bantu languages predominate, was a major embarkation region in the first quarter of the eighteenth century, after which the numbers dropped to zero, bouncing back only in the 1790s. The captives who landed in

⁷This text is traditionally attributed to the plantation owner J. L. Carstens (1705–1747), but there is "plenty of evidence that Carstens [...] did not write it," as Sebro (2016: 219) cautions.

⁸"Other Africa" is for the most part used for captives whose point of embarkation is unknown, i.e., 31,052 (35%) of the people accounted for in Table 1 (the remaining 1,150 people of the 32,220 subsumed under this category were tracked to more specific locations). The category is in practice often used for voyages that acquired captives along the Windward Coast. This required several stops, as only small numbers of individuals could be negotiated at the different ports of call (cf. Vos 2010).

2 African ethnolinguistic diversity in the colonial Caribbean

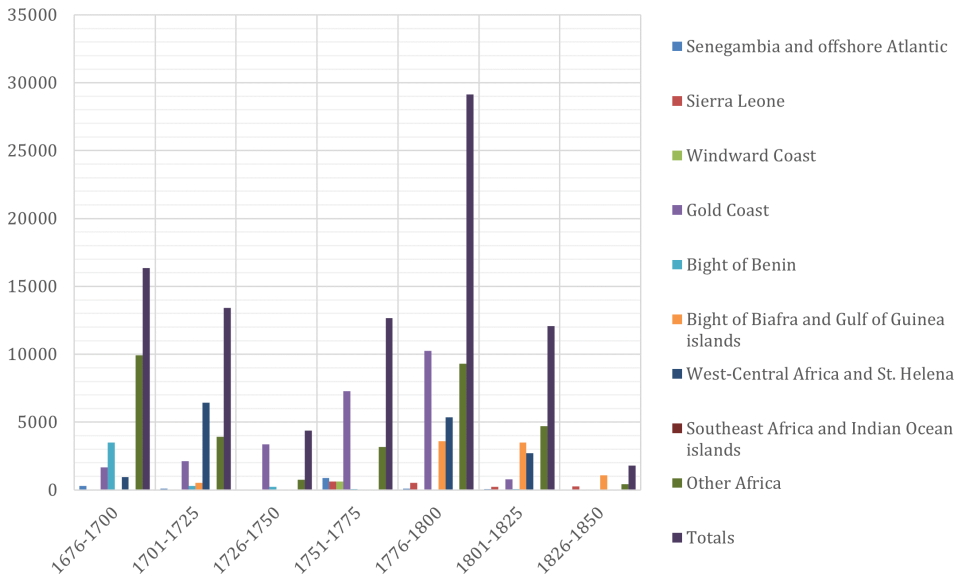


Figure 1: Proportions of enslaved people from the different embarkation regions

the later period came from a more diverse set of ethnic backgrounds, reflecting the fact that the slave trade spread to a wider range of regions in the latter part of the eighteenth century. The Gold Coast still supplied the most enslaved people to the colony, whereas the Bight of Benin evidently was of much less relevance in this later phase. Moreover, between the mid-1770s and 1800, the Bight of Biafra, which is home to a different group of languages (especially different Benue-Congo subgroups), contributed about 12% of the enslaved people landed in the Danish West Indies. In addition, Senegambia, Sierra Leone, and the Windward Coast, which likewise are representative of a different range of languages (Mande, North-Atlantic, Kru, etc.), continuously supplied smaller quantities of human cargo, which, as noted, can be presumed to have been larger in reality than the recorded trade indicates.

In evaluation, it would appear plausible, based on these data, that “substantial ‘ethnic diversification’ did not occur until the latter part of the eighteenth century,” as Sabino (2012: 66) has argued. However, the transatlantic trade was not the only source of slave imports to the islands. Indeed, when factoring in other sources of evidence, it becomes difficult to uphold what might be termed the “heavy Kwa bias” hypothesis, at least in its strong form. The remainder of the study corroborates this view.

Table 1: Estimates of slave imports to the Danish West Indies via the transatlantic trade in 25-year periods, arranged by embarkation regions (based on the Voyages Database, accessed 30 September 2019). Absolute numbers in plain font, percentages for each period in italics.

Year range	Senegambia and offshore Atlantic	Sierra Leone	Windward Coast	Gold Coast	Bight of Benin	Bight of Biafra and Gulf of Guinea islands	West-Central Africa and St. Helena	Southeast Africa and Indian Ocean islands	Other Africa	Totals
1676–1700	299	0	0	1,658	3,508	0	938	0	9,937	16,340
<i>Percentage</i>	<i>1.8</i>	<i>0.0</i>	<i>0.0</i>	<i>10.1</i>	<i>21.5</i>	<i>0.0</i>	<i>5.7</i>	<i>0.0</i>	<i>60.8</i>	<i>100.0</i>
1701–1725	97	0	0	2,124	295	536	6,433	0	3,918	13,403
<i>Percentage</i>	<i>0.7</i>	<i>0.0</i>	<i>0.0</i>	<i>15.8</i>	<i>2.2</i>	<i>4.0</i>	<i>48.0</i>	<i>0.0</i>	<i>29.2</i>	<i>100.0</i>
1726–1750	0	0	0	3,368	242	0	0	21	742	4,373
<i>Percentage</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>77.0</i>	<i>5.5</i>	<i>0.0</i>	<i>0.0</i>	<i>0.5</i>	<i>17.0</i>	<i>100.0</i>
1751–1775	883	639	636	7,284	62	0	0	0	3,172	12,676
<i>Percentage</i>	<i>7.0</i>	<i>5.0</i>	<i>5.0</i>	<i>57.5</i>	<i>0.5</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>25.0</i>	<i>100.0</i>
1776–1800	99	543	0	10,251	0	3,602	5,346	0	9,312	29,153
<i>Percentage</i>	<i>0.3</i>	<i>1.9</i>	<i>0.0</i>	<i>35.2</i>	<i>0.0</i>	<i>12.4</i>	<i>18.3</i>	<i>0.0</i>	<i>31.9</i>	<i>100.0</i>
1801–1825	58	220	53	780	61	3,485	2,724	0	4,695	12,076
<i>Percentage</i>	<i>0.5</i>	<i>1.8</i>	<i>0.4</i>	<i>6.5</i>	<i>0.5</i>	<i>28.9</i>	<i>22.6</i>	<i>0.0</i>	<i>38.9</i>	<i>100.0</i>
1826–1850	0	277	0	0	0	1,087	0	0	444	1,808
<i>Percentage</i>	<i>0.0</i>	<i>15.3</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>60.1</i>	<i>0.0</i>	<i>0.0</i>	<i>24.6</i>	<i>100.0</i>
Totals	1,436	1,679	689	25,465	4,168	8,710	15,441	21	32,220	89,829
<i>Percentage</i>	<i>1.6</i>	<i>1.9</i>	<i>0.8</i>	<i>28.3</i>	<i>4.6</i>	<i>9.7</i>	<i>17.2</i>	<i>0.0</i>	<i>35.9</i>	<i>100.0</i>

3.2 The intra-Caribbean slave trade

In the eighteenth century, Simonsen & Olsen (2017: 150) note, “[i]t was the import of enslaved people from Africa and, secondarily, the import of enslaved people from the surrounding Caribbean islands, inter alia St. Christopher [i.e., St. Kitts, KFB], Anguilla, and Tortola, which created the population growth in the Danish West Indies” (my translation). I will not attempt to account for the local trade in its totality, part of which is unrecorded and hence unquantifiable, but I will offer an example that can illustrate what it implies in terms of modifying and nuancing the picture based solely on the transatlantic slave trade data.

Between 1766 and 1802, based on the years for which reliable figures are available (cf. Green-Pedersen 1975: 208–209), approximately 30,150 enslaved people were imported to St. Croix, both directly from Africa and via the local intra-Caribbean trade. Of these people, 8,444 were landed via the local trade, i.e., at least some 28% of the total number. As these figures do not take into account the unrecorded and illicit trade (cf. Gøbel 2016: 174 and the references therein), the actual share was larger. Moreover, when looking into individual years, this percentage was often higher yet. This was especially so prior to the ten-year grace period following 1792, the year that the proclamation abolishing the Danish slave trade ten years hence went into effect, which prompted a significant increase in the transatlantic trade to the Danish West Indies (cf. Gøbel 2016: 138 ff.).

Based on inspection of tax rolls for the period 1767–1776 in the Danish National Archives, Olsen (1988: 229) calculated the percentages of enslaved people landed on St. Croix from Africa and other (i.e., primarily, Caribbean) locations during this period. Table 2 summarizes these results. For a number of the years accounted for by Olsen, the majority of the enslaved landed via the recorded trade did not actually arrive at the islands on vessels directly from Africa. Over the entire period, roughly half of the enslaved came from other parts of the Atlantic world, many of them from other Caribbean islands. It should be noted that some of the enslaved landed via this trade did, in fact, come from Africa but had been re-exported via other islands (cf. Olsen 1988: 233–235). Needless to say, this either complicates the picture or renders impossible the task of tracing them to specific embarkation points.

Due to the extensive local trade, it follows that it is necessary, in addition, to consider other sources of evidence for the origins of the enslaved population in the Danish West Indies. Accordingly, the remainder of the study will focus on the language and ethnicity documentation provided by the Moravian Brethren, recorded locally in the colony.

Table 2: Percentages of enslaved people landed on St. Croix from Africa and elsewhere, 1767–1776 (based on Olsen 1988: 229)

Year	Slave imports	Directly from Africa	
		N	%
1767	1,962	468	23.8
1768	1,686	1,163	70.1
1769	–	(68)	–
1770	1,862	1,131	60.7
1771	1,157	835	72.1
1772	850	588	69.1
1773	241	0	0.0
1774	1,017	321	31.5
1775	1,670	743	44.4
1776	1,080	706	65.3
1767–1776	11,525	6,023/5,955	52.0

4 Oldendorp’s interview data

“Ideally,” Kouwenberg (2008: 2) states, the substratist researcher “hope[s] to find clear and incontrovertible linguistic evidence” for the presence of particular African languages, “supported by the historical record.” Following this lead, this section aims to establish credible links between the African languages documented in the historical material from the Danish West Indies with modern-day correlate languages and language groups. To this end, I analyze linguistic data and (other) cultural information recorded by C.G.A. Oldendorp (1721–1787) of the Moravian Brethren for an ethnographic survey.

Oldendorp spent about two years in the Danish West Indies, from 1767 to 1768. His stay resulted in a highly influential account of the islands and the Moravian Mission’s history up to that point in time, including a detailed description of the life and cultural affinities of the enslaved in the colony (for background information on Oldendorp’s account, see Meier et al. 2010).⁹ Oldendorp completed his massive report in 1776, and it was published – albeit in significantly abridged form – the following year (Oldendorp 1777).¹⁰ In 1987, an English translation of

⁹In addition, Oldendorp also compiled his *Criolisches Wörterbuch* (1996) during his stay.

¹⁰For a discussion of Oldendorp’s reaction to the abridgment of his report, see, e.g., Ahlbäck (2016).

the 1777 edition appeared (Highfield & Barac 1987), which was then followed by a critically edited multi-volume version of the original manuscript in its entirety (Oldendorp 2000, 2002). Note that in the following, I quote from both the critically edited German version and the English translation of the account.

An essential part of Oldendorp's account is an ethnographic survey, based on interviews with approximately 70 informants in St. Thomas and St. Croix. Oldendorp's study was to provide insights into the conditions of the enslaved by mapping the various "nations" represented in the colony, and by providing an overview of the various religious ideas represented among them. Oldendorp was well aware that there were cultural differences among the enslaved, remarking that "[t]he Negroes are divided into many nations, among which greater or lesser differences can be found in regard to language, as well as morals, customs, and religion" (Highfield & Barac 1987: 160). Oldendorp describes how he gathered his informants into peer groups according to their diverse origins, and how he then asked them about their (ancestral) religious customs and practices, the geographical location of their homelands, neighboring groups (both allies and enemies), trade relations, etc. There is reason to highlight the importance of the fact that Oldendorp based his description on information from groups that agreed that they belonged to the same ethnic category, and on that basis shared information. When several people agreed to accept each other as peers, there is good reason to believe that the information that came out of it also reflected real notions of a community, or at least cultural commonalities (cf. also Sebro 2010: 69).

In addition, Oldendorp collected linguistic data from his interviewees. He recorded label-meaning correlations of the numerals 'one' to 'ten' (and selected higher numbers) across 22 languages (cf. Oldendorp 2000: 458–460), as well as a set of thirteen "basic" meanings, such as, e.g., 'head' and 'sun,' across 26 languages (pp. 460–463), with some gaps in the distribution. There is, in addition, a single sentence translated into the various languages: *Christus hat uns geliebet und gewaschen von den Sünden mit seinem Blut* 'Christ has loved us and has washed away our sins with his blood' (pp. 464–465). According to Jones (2010: 181), the full corpus consists of 667 words in the different African languages.

Ahlbäck (2016: 194) notes that the publication of Oldendorp's original manuscript in 2000–2002 "has been welcomed as providing an outstanding source to eighteenth-century African Caribbean cultures, languages and identities." As of yet, however, no actual language-by-language examination of Oldendorp's linguistic data, as found in this version of the account, has appeared. The two previous analyses, namely Peter Stein's editorial notes in Oldendorp (2000) and Jones' (2010) useful overview, constitute valuable contributions to such an end, as does the examination by Fodor (1975) of the 1777 edition. In the following, I consider,

in a concise analysis, the information collected by Oldendorp, identifying the languages (or, alternatively, the closest matches found) with present-day correlates organized by affiliation, on the basis of the linguistic data. I compare Oldendorp's collected numerals from 'one' to 'ten' with precisely transcribed modern lists of numerals in African languages. Unless otherwise noted, the modern numerals have been extracted from the database *Numeral Systems of the World's Languages* – an online repository of basic data on more than 4,000 of the world's languages, formerly hosted at the Department of Linguistics at the MPI for Evolutionary Anthropology in Leipzig.¹¹ Where available and where relevant (i.e., in practice, where the numerals data are either lacking or inconclusive), I refer also to some of the other linguistic data recorded by Oldendorp. In addition, I refer to cultural and geographic clues offered in the account, as well as to previous work within the Africanist literature and the aforementioned previous analyses of Oldendorp's material.

4.1 Identification of languages in Oldendorp's account

The sets of numerals referred to in the following analysis are presented collectively in Table 3. The presentation of languages is based on the order in which they are discussed below. All modern languages are identified by their three-letter ISO 639-3 codes, marked by square brackets. The designations used by Oldendorp (which may represent ethnonyms or glossonyms, or both) are presented within double quotation marks. All languages which Oldendorp recorded direct evidence on are represented in the discussion.

4.1.1 “Fula”: Fula, North Atlantic

Oldendorp's “Fula” numerals match those of varieties in the Fula cluster [inclusive ISO code: ful], except for the forms for ‘eight’ and ‘nine.’ This discrepancy is most likely due to Oldendorp, or his informant, having switched the forms *jenai* and *jädet* around. In addition, the “Fula” informant reported having traveled for two months to reach the ocean from his homeland, located along a great river (Oldendorp 2000: 373), which was probably the Senegal or Gambia River. Pulaar [fuc] is included in Table 3 as the probably closest-matching candidate lect.

¹¹The database is currently available at <https://lingweb.eva.mpg.de/channumerals/> (last accessed 29 November 2024). The database forms the basis for the “Numeralbank” project carried out under the Glottobank research consortium: <https://glottobank.org/> (last accessed 29 November 2024).

2 African ethnolinguistic diversity in the colonial Caribbean

Table 3: Numerals from ‘one’ to ‘ten’ across 23 languages recorded by Oldendorp compared with data on 26 modern African (Niger-Congo) languages and their subgroups

“Nation”/language	1	2	3	4	5	6	7	8	9	10
North Atlantic (region: Senegambia)										
“Fula” Pulaar [fuc]	go goo	diidi diidi	taddi tati	nei naj	djoi dʒoj	djogo dʒegom	tjedidi dʒeedidi	jenai dʒeetati	jädet dʒeenaj	sappoi sappo
Mande (region: interior Windward Coast)										
“Jalunkan” Yalunka [yal]	keling kedé	filla fir-ŋ	saba sakaŋ	nani nani	lolu sulu	worro seni	orwila foloféré	sagi folomāsakán	kononto folómánani	tan fú
Dyula [dyu]	kelen	filá	sábá	nánin	diúrú	wòrò	wolon filá	siegi	kónánón	tan
Jakunan [bxl]	dáli	filá	sigbù	nāni	sōlō	mi:lù	mādlá	māsīgḅū	mānāni	tā
“Sokko”	külle	felaa	sauaa	nani	duli	woro	ornala	setti	konundo	tang
Gur (region: North Ghana, South Burkina Faso)										
“Tembu” Tem [kdh]	kuddum kááde	noalee sié	nodosoo tóozo	nonasaa náaza	nonoaa néówa	lodo loqo	lubbe lobé	lütoso lutoozo	kandilee kéeníré	figuh fúú
Lukpa [dop]	kèlém	naalé	tóosó	naasá	kákpásì	naátosó	naátosóm-py:ɔlayá	pálé fėjí	py:ɔlayáfėjí	náaná
“Tjamba” Kasem [xsm]	obaa kalí	illee nlé	itaa nlé	inna hna	immu hnu	iloop hdon	illetee hpe	imenn nana	üwáh négó	piek fúgá
Konkomba [xon]	-báa	-lée	-táa	-náa	-nmúu	-lúub	-lilé	-niin	-wéé	pük
Kru (region: Liberia, Ivory Coast)										
“Gien” Krahn [krw]	do tò	sung sòn	ta táan	nje nyè	mu mih	meu mèò	mesong mésòon	medda metáàn	menje mèryiè	wo pièè
“Kanga” Kyanga [tye]	aniandu diú	aniasson f’áa	anietan ’áa:	anaje fi	aneamu sòrú	— péniú	— ji	— sáqós	— dókó	anièpun íi

Table 3: Numerals (continued)

"Nation"/language	1	2	3	4	5	6	7	8	9	10
Gbe, Kwa; Yoruboid, Benue-Congo (region: Togo, Benin)										
"Papaa"	dépoò	auwi	ottong	enne	attong	attugo	atjuwe	attiatong	atjenne	awo
"Wavu"	dépoò	aweé	etong	enne	attong	a-isee	djauí	tiatong	tienee	wó
Anders (or: "Wavu II")	baba	batili	Janna	toffa	guena	brong	Je-gra	khiboa	boafri	magro
"Wajje"	de	e-wee	etong	enne	attong	andee	anderee	enni	enjidee	owoo
Fon [fon]	dǝ	wɛ	atɔn	enɛ	atɔ́n	ayizén	ténwe	tánton	ténne	wò
Aja [ajig]	edǝ, dǝka	evè, amɛ'vɛ	etǝ, amɛitǝ	enɛ, amɛnɛ	atǝ, amaitǝ	adɛ, amaidɛ	adɛ, amaidɛ	ɛpi, amɛpi	jidɛ, ajidɛkɛ	ewó
Ewe [ewe]	edǝ	evè	étǝ	ènè	átǝ	àdè	adrè	ɛpi	afieké	ɛwó
Gã-Dangme, Kwa (region: coastal Ghana)										
"Tambi"	kaki	ennu	ette	ewe	enu	ekba	pagu	panjo	ne	njomma
"Akkran"	eaku	eenjo	ette	eedjee	ennumo	epa	paggu	paniu	nehung	jungma
Gã [gaa]	ékómé	éjɔ̀	étɛ	éjwè	énúmɔ̀	ékpáa	kpáwo	kpáapɔ̀	nɛchú	jònmá
Dangme [ada]	kakɛ	éjɔ̀	étɛ	éjwè, éwiè	énúɔ̀	ékpá	kpáagò	kpáapɔ̀	nɛɛ́	jònmá
Cross-River, Benue-Congo; Igboid, Benue-Congo (region: Nigeria)										
"Karabari"	otuh	abolam	attoo	abanna	abisee	abisih	abassa	abassatto	abitolu	abilli
"Ibo"	otuh	aboa	attoo	anoo	issee	issih, fschi	assaa	assatto	ilelite	ili
Ijo Kalabari [ijn]	ɔ̀jɛ̀i	mǎ	terɛ	ineí	sonɔ	sonio	sonomè	ninè	esenie	oji, àtèi
Igbo [ibo]	otú	ábɔ̀ɔ̀	átɔ̀	ánɔ̀	isè	Isi	ásátɔ̀	ásátɔ̀	ifolu	iri
"Mokko"	kiá	iba	itta	inan	úttin	itúiekee	ittiaba	ititeata	huschukiet	büü
Efik [efi]	kiét	ibá	itá	ináj	itíon	itíókíét	itábá	itáitá	úsuk-kiét	diúop
Bantu, Benue-Congo (region: coastal and interior West-Central Africa)										
"Loango"	bosse	quari	tattu	ena	tanu	sambaan	sambuari	nane	iwoa	kumi
"Congo"	moschi	sole	sitattu	sija	sittan	issamban	samboari	sinaan	siwoa	sikumi
"Camba"	moschi	soli	tattu	ja	tanu	saman	sambari	nane	wa	komi
Kikongo [kng]	mási	zòlè	tátù	yá	tánù	sámánù	nsámwáadí	nàná	vvá	kùmi
"Mandongo"	omma	meere	metutu	mína	metaan	schiavano	entschewine	ennane	woa	kumi
Kimbundu [kmb]	móxi	yádi	tátù	wáná	tánù	sámánù	sámwáadí	(d)vwá	(d)vwá	(d)kwinyi
Ndumu [nmd]	-mɔ	-eɛɛ, -ɔɔɔ	-tati	-na	-taani	-samani	tsaami	pwɔmɔ	wua	kumu
Mbere [mbt]	-mɔ	-ele	-fare	-na	-taani	-syaami	nisaami	mpfwamɔ	wa	kuomi

4.1.2 “Mandinga”: Mandinka, Mande

Oldendorp described the “Mandinga” (or “Mandingo”) as a neighboring tribe of the “Fula” and “Jalunkan” (see below), a people related to them but differing in language (Oldendorp 2000: 376). Oldendorp recorded no “Mandinga” numerals, but he did include word forms for some of the other basic meanings. Table 4 shows that correspondences with the Mande language Mandinka [mnk] can be found for four of six of these (Mandinka data adapted from Peace Corps The Gambia 1995). Thus, Oldendorp’s “Mandinga” most likely corresponds to a Mandinka lect, but the linguistic evidence is too scanty to allow any precise identification.

Table 4: Oldendorp’s “Mandinga” compared with Mandinka [mnk]

	‘God’	‘Sun’	‘Mouth’	‘Hand’	‘Father’	‘Mother’
“Mandinga”	<i>Allah, Kanniba</i>	<i>tille</i>	<i>pandintee</i>	<i>bullla</i>	<i>ba</i>	<i>jem</i>
Mandinka [mnk]	<i>Ala</i>	<i>tili</i>	<i>dáa</i>	<i>búlu</i>	<i>baaba</i>	<i>baa, naa</i>

4.1.3 “Jalunkan” and “Sokko”: Dyula-Bambara, Mande

It is suggested by Peter Stein (Oldendorp 2000: 376, fn. 54) that Oldendorp’s “Jalunkan” refers to the Yalunka language [yal] of Guinea and Sierra Leone. Conversely, Fodor (1975: 104) suggested that the language could be Dyula [dyu], which is mutually intelligible with the West African lingua franca (and, since the 1960s, a national language of Mali) Bambara [bam]. In addition, “Jalunkan” could also refer to Jalkunan [bxl] of Burkina Faso. Note that the modern Yalunka numerals ‘seven’ through ‘nine’ are compounds based on a unit ‘five.’ The collected numerals display closest alignment with those of modern Dyula.

Oldendorp noted about the “Sokko” (or “Asokko”) that they were Islamic, and that it took them from six to seven weeks to reach the coast from their home region, neighboring the “Amina” (Oldendorp 2000: 407–408), i.e., peoples of the Gold Coast (cf. Law 2005: 247–248). “Sokko” may well refer to the Dyula town Begho-Nsoko in interior Ghana, close to the Ivory Coast border (Jones 2010: 186; Oldendorp 2000: 407, fn. 84; Stahl 2001: 124). Migeod (1913) linked the area’s people, whom Koelle (1854) called “Jalonke,” with the Dyula, and he wrote that they self-identified as the “Sako” (Migeod 1913: 346). Indeed, the “Sokko” numerals recorded by Oldendorp show close similarity to those of Dyula-Bambara. Moreover, nine out of thirteen of the “Jalunkan” and “Sokko” lexical items in Oldendorp (2000: 460–461) show close similarity, corroborating further that also the “Sokko” used a Dyula-Bambara lect.

4.1.4 “Tembu”: Tem, Gur

The “Tembu” were noted to live further inland than the “Amina,” four days’ travel from the land of the coastal “Akkran” (Oldendorp 2000: 399–400). The “Tembu” numerals match those of the interior Gur language Tem [kdh] closely, except for the form for ‘two.’ The related language Lukpa [dop] has a form corresponding to that for ‘two’ given by Oldendorp, but this language does not otherwise match the “Tembu” numerals closely. Overall, the “Tembu” data recorded by Oldendorp are closest to Tem.

4.1.5 “Tjamba”: Konkomba, Gurma, Gur

Oldendorp reported that his “Tjamba” informants had traveled for as long as six months to reach the Gold Coast from their homeland, and that their king had his residence in the city of “Gambaak” (Oldendorp 2000: 403–404), i.e., Gambaga in the North East Region of Ghana. Jones (2010: 186) has suggested that Oldendorp’s “Tjamba” (also called “Kassenti” – supposedly meaning ‘I do not understand you’) refers to the Kasem, or Chamba, language [xsm] of southern Burkina Faso, a language of the Gur group, but the linguistic data do not lend support to this view. Instead, the “Tjamba” numerals match those of Konkomba [xon] of the Gurma cluster, also a Gur language, spoken in Northern Ghana.

4.1.6 “Gien” and “Kanga”: Kru lects

According to Westermann & Bryan (1952: 50–51), the name “Gien” was used in reference to the Southern Wee language Krahn [krw] of the Ivory Coast and Liberia. Indeed, the collected numerals match closely between Oldendorp’s “Gien” and Krahn. Another candidate language is the closely related lect Grebo [grj] (Grebo numerals adapted from Innes 1967). As for the “Kanga,” although there is a Mande language called Kyanga [tye], the numerals (of which Oldendorp did not record ‘six’ to ‘nine’) indicate it was a Kru lect, too. Note that none of the modern Kru lects considered retains the (presumed) prefix *ania-*, or *anie-*, included by Oldendorp. Some additional clues given by Oldendorp indicate that these languages were varieties of Kru, and not, e.g., Mande. Oldendorp indicated that the ocean formed the western border of the “Kanga” territory, and that they often traded with Europeans. With their land bounded to the west by the sea, they otherwise shared borders with the “Mandinga” and “Fula” – but they did not understand their languages (Oldendorp 2000: 378). Mandinga and Fula are Mande and North Atlantic, respectively, indeed very different from Kru languages.

4.1.7 “Mangree”: Another Kru language?

Oldendorp (2000: 382) noted that a great river, which must be either the Sasandra or the Bandama (cf. Fodor 1975: 162), constituted the border between the “Gien” and the “Mangree,” and that their languages “[did] not differ much from one another” (Highfield & Barac 1987: 162). Oldendorp recorded no “Mangree” numerals (for which reason the language does not figure in Table 3), but he noted that the “Mangree” people understood the “Kanga” (2000: 378), and that they lived deep in the interior. In addition to the “Gien,” the “Mangree” lived near the “Mandinga” and “Amina” (2000: 381–82). Thus, the “Mangree” could correspond to the Ngere (or Wee) of the Ivory Coast (cf. Jones 2010: 186; Oldendorp 2000: 381, fn. 58) who spoke a Wee lect (close to that of the Krahn-speaking “Gien”). Oldendorp did record some basic “Mangree” lexicon. There are few correspondences, however, between the “Mangree” and, respectively, “Gien” and “Kanga” items (cf. Oldendorp 2000: 460–461).

4.1.8 “Amina” and “Akkim”: Akan, Potou-Tano, Kwa

The “Amina” were described by Oldendorp as the most powerful “nation” east of the Gold Coast. Their land extended from close to the coast and well into the interior, and they waged war on their various neighbors (Oldendorp 2000: 383). The “Amina” numerals recorded by Oldendorp correspond closely to those of the Akan [aka] lects Twi and Fante (Fante numerals adapted from Bureau of Ghana Languages 1986), except for the form for ‘two’ – and in the case of Fante, ‘one’. Moreover, the language of the “Amina” was described as being the same as that of the “Akkim” tribe (corresponding to modern-day Akyem), which is corroborated by the linguistic data. Oldendorp (2000: 392) noted about the “Akkim” that they lived close to the sea, a day’s journey from Danish Fort Christiansborg. For a recent in-depth study on Oldendorp’s “Amina,” see also Kelley & Lovejoy (2023).

4.1.9 “Okwa”

In addition, Oldendorp (2000: 396) noted having spoken with one individual from the same region as the two aforementioned groups who self-identified as an “Okwa” (or “Okwoi”). However, as only the word *Tschabee* ‘God’ was recorded in this language, it has not been possible to establish its identity.

4.1.10 “Akripon”: Larteh, South Guang, Potou-Tano, Kwa

Oldendorp (2000: 395) noted that the “Akripon” shared a common border with the “Amina,” but they constituted a separate kingdom, and they also used the

same language. “Akripon” is most likely a toponym corresponding to the modern town of Akropon in southern Ghana. The South Guang language Larteh [lar] is spoken there (Bello 2013). Its numerals match Oldendorp’s “Akripon” numerals closely. Some North Guang lects match for the most part, but forms resembling Oldendorp’s *ebnoo* ‘nine’ are only found in the South Guang cluster.

4.1.11 “Papaa” and “Wavu”: Fon, Gbe, Kwa

Both the “Papaa” and “Wavu” numerals match those of modern Fon [fon] (Fon numerals extracted from Lefebvre & Brousseau 2002). Some other differences can, however, be noted between the two seemingly distinct groups. Oldendorp (2000: 411–412) stated that the “Affong” were the rulers of the “Papaa” kingdom, and that his “Papaa” informants were familiar with both the Danes and other Europeans, as well as with the “Akkran” and “Amina” (from whose raids they suffered). Their land extended to the sea on one side. The “Wavu” likewise lived in part on the coast, but also in part deep within the interior where they comprised a populous nation “throughout which the same language is not spoken uniformly” (Highfield & Barac 1987: 166).

4.1.12 “Wavu II”

Fodor (1975: 132–137) presented an in-depth discussion of Oldendorp’s “Wavu,” which he argued actually referred to two distinct languages. The first of these is the Gbe variety identified above; the second – labeled *anders* ‘different’ (i.e., different from “Wavu”) by Oldendorp and “Wavu II” by Fodor – remains obscure.¹² I have not been able to link “Wavu II” to any modern language.

4.1.13 “Watje” and “Atje”: Aja/Ewe, Gbe, Kwa

Oldendorp (2000: 413, 419) noted that the “Watje” were the rulers of a separate kingdom, their territory extending far inland from the sea. Their neighbors included the “Amina” (whom they were at war with), the “Tjamba” (or “Kassenti”), and the “Sokko.” The “Watje” and “Atje” were described as closely related groups, with languages that were “almost identical” (Oldendorp in Highfield & Barac 1987: 166). Oldendorp recorded no “Atje” numerals, but the modern Gbe lects Aja [ajg] and Ewe [ewe] match the “Watje” numerals closely.

¹²Cf. the entry for “Wavu II” in *Glottolog*: <https://glottolog.org/resource/languoid/id/wavu1234> (last accessed 29 November 2024).

4.1.14 “Tambi” and “Akkran”: Gã-Dangme, Kwa

The “Tambi” and “Akkran” groups were described as living next to the Danish fort. They were reported to understand the language of the neighboring war-like “Amina,” although they did have a language of their own (Oldendorp 2000: 396, 399, 411). The “Akkran” and “Tambi” numerals show alignment with the two closely related Kwa languages Gã [gaa] and Dangme [ada], spoken in and near Accra, Ghana.

4.1.15 “Karabari” and “Ibo”: Igbo, Benue-Congo

Oldendorp (2000: 426) described the “Karabari” as living on the Calabar River, far from the sea. The “Ibo,” a numerous people who lived in a vast land in the interior, were their neighbors, and were noted to use the same language. Peter Stein (Oldendorp 2000: 426, fn. 108) has suggested that the “Karabari” were speakers of Ijo Kalabari [ijn] (Ijoid), but the collected numerals resemble those of modern Igbo [ibo], not any of the Ijoid languages. Indeed, “Karabari” was a generic term used about peoples living along the Calabar River (cf. Fodor 1975: 138). Also the “Ibo” numerals match those of modern Igbo.

4.1.16 “Mokko”: Efik, Cross-River, Benue-Congo

The “Mokko” shared a border with the “Karabari” (Oldendorp 2000: 434). The “Mokko” designation is not unproblematic, in that it was used for several peoples (Oldendorp 2000: 434, fn. 115), but in this context it most likely refers to the Efik-speaking Ibibio of Nigeria (Hair 1967: 263). Indeed, the “Mokko” numerals recorded by Oldendorp correspond to those of modern-day Efik [efi].

4.1.17 “Loango,” “Congo,” and “Camba”: Kikongo, Bantu, Benue-Congo

Oldendorp (2000: 436–445) noted that the “Loango” lived about a month’s journey from the West-Central African coast, which can presumably be identified with present-day Gabon. The “Congo” lived close to the coast, and one of them knew the Portuguese-founded city Luanda, the capital of present-day Angola. The “Camba” lived near the “Loango,” not far from the “Sundi” (or “Nsundi”). According to Dalby (1964: 85), the group described as “Sundi” by Koelle (1854) lived in Guthrie’s (1948) Bantu Zone H. The numerals of Oldendorp’s “Camba” and “Congo” match those of Kikongo [kng], except that Oldendorp’s “Congo” takes what may be a concord prefix, *si-*, for ‘three’ to ‘six’ and ‘eight’ to ‘ten.’ The “Loango” numerals are similar to those of Kikongo, with the exception of

the “Loango” form *quari* ‘two.’ Moreover, eleven out of thirteen of the “Loango” and “Congo” lexical items show close similarity, and the same is the case for ten out of thirteen “Camba” and “Loango” items. Thus, the three groups all spoke varieties of Kikongo.

4.1.18 “Mandongo”: A language of Bantu Zone B or H

Finally, Oldendorp described the “Mandongo” as a widely dispersed people consisting of three scattered groups, “bound to one another by a common language” (Highfield & Barac 1987: 168). Some of Oldendorp’s “Mandongo” interviewees indicated that it took an entire year for the journey from their homeland to the land of the “Loango,” and from there it took approximately one month to reach the sea (Oldendorp 2000: 442). It is suggested by Peter Stein (Oldendorp 2000: 441, fn. 123, citing Birmingham 1966) that the “Mandongo” designation corresponds to Ndongo, a Mbundu chiefdom near Luanda, in Guthrie’s (1948) Bantu Zone H (roughly, northwestern Angola to western Congo). However, there is only partial alignment between Oldendorp’s “Mandongo” and the then expected correlate language Kimbundu [kmb]. Fodor (1975: 152) suggested that the “Mandongo” language belonged in the Mbete (or Mbere) cluster of Bantu Zone B (roughly, southern Gabon over western Congo), the same as, e.g., Ndumu [nmd] or Mbere [mdt]. Its precise identity remains uncertain, however. The collected numerals show closest similarity with various Zone B and H Bantu languages, but no one modern Bantu language emerges as a surefire match.

4.2 Summary and short assessment

Whereas early African language materials (where available) often suffer from imprecise transcription (Jones 1991: 95–97), Oldendorp’s reporting has long been recognized, and commended, for its accuracy (e.g., Herskovits 1958: 44; see also, e.g., Ahlbäck 2016). Corroborating this view, the above analysis has shown that, for the most part, the linguistic data in Oldendorp’s account are readily comparable with data on modern African languages. Differences could be due to dialectal varieties, or language changes in the past two and a half centuries. Oldendorp’s “Mangree” and “Okwa” could not be identified with any present-day language, neither could the language Fodor (1975) labeled “Wavu II.” These cannot be ruled out to have been spurious. Further, the Bantu language “Mandongo” was only tracked to an approximate location on the West-Central African coast. Besides these, however, it was possible to link the remainder of Oldendorp’s languages confidently with modern-day correlates (as summarized in Table 5).

2 African ethnolinguistic diversity in the colonial Caribbean

Table 5: Modern-day correlates to the African languages documented by Oldendorp

Language in Oldendorp's account	Modern-day African language(s)	Affiliation
"Fula"	A language of the Fula cluster [ful], perhaps Pulaar [fuc]	North Atlantic
"Mandinga"	Presumably Mandinka [mnk]	Mande
"Jalunkan" and "Sokko"	Dyula-Bambara [dyu]	Mande
"Tembu"	Tem [kdh]	Gur
"Tjamba"	Konkomba [xon]	Gur
"Gien" and "Kanga"	Kru lects, presumably Krahn [krw] and/or Grebo [grj]	Kru
"Mangree"	–	Kru, presumably
"Amina" and "Akkim"	Akan [aka] lects	Kwa
"Okwa"	–	–
"Akripon"	Larteh [lar]	Kwa
"Papaa" and "Wavu"	Fon [fon], Gbe	Kwa
"Watje" and "Atje"	Aja [ajg] and/or Ewe [ewe], Gbe	Kwa
"Tambi" and "Akkran"	Gã [gaa] and/or Dangme [ada]	Kwa
"Karabari" and "Ibo"	Igbo [ibo], Igboid	Benue-Congo
"Mokko"	Efik [efi], Cross-River	Benue-Congo
"Loango," "Congo," and "Camba"	Kikongo [kng], Bantu	Benue-Congo
"Mandongo"	A Bantu language of Zone B or H	Benue-Congo

The map in Figure 2 shows the approximate geographic distribution of the African languages found to be represented in the Danish West Indies based on the analysis of Oldendorp's material. In line with the analysis in §3, the languages can be seen to come from a stretch that extends from present-day Senegal to Angola – over more than 4,000 km. Most languages are spoken off the coastline from Ghana to Benin, but also Nigerian and West-Central African languages have a clear representation, as do languages from the continent's westernmost coast. Strikingly, several languages are from far into the interior, months of travel from the coast.

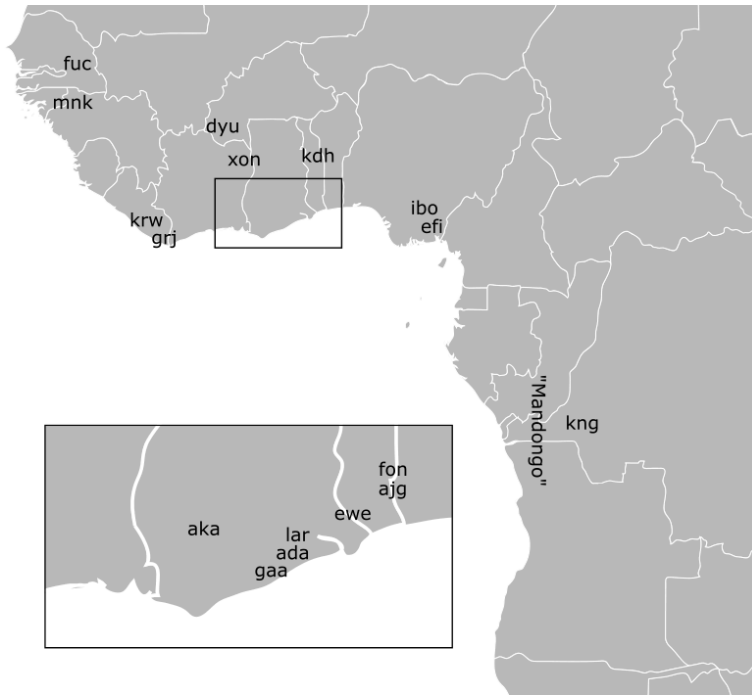


Figure 2: Approximate locations of the identifiable languages listed in Table 5, represented by ISO 639-3 codes (excepting the imprecisely identified “Mandongo” Bantu language)

Building on the above analysis, the following two sections will examine Seidel’s list of Christianized Africans and African Caribbeans from St. Thomas and the baptismal records from St. Croix, looking at similarities and differences in the distribution of the identified ethnolinguistic groups and their proportions between the two islands.

5 St. Thomas: Seidel’s list of Christianized Africans and African Caribbeans

In 1753, the Moravian missionary Nathanael Seidel (1718–1782) visited the Danish West Indies for a duration of two months. During his time there, he recorded observations for a report on the status of the mission work (Seidel 1753). As part of his report, Seidel compiled a list of the different African “nations” represented among the communicants of the Moravian congregation on St. Thomas,

2 African ethnolinguistic diversity in the colonial Caribbean

i.e., those of the enslaved who were entitled to receive communion. The report was unearthed from the Moravian Archives in Herrnhut, Germany, by the Danish historian Louise Sebro, in connection with her archival research for Sebro (2010), in which the information pertaining to the interviewees' various origins was tabulated as an appendix (pp. 215–16). Seidel interviewed 418 communicants in St. Thomas, representing some 12% of the island's total enslaved population at the time (estimated on the basis of data from the Danish National Archives in Green-Pedersen 1971: 157). All but 36 (8.6%) of Seidel's informants stated ethnic groups/places of origin, which are potentially identifiable and indicative of the range of ethnolinguistic groups represented in St. Thomas. Expanding on Sebro's work on Seidel's list, the present analysis is the first to make its contents available to an international audience (although see Bakker 2016b: 226 for a brief characterization).

Along with Oldendorp's interview data (§4) and the baptismal records from St. Croix (§6), Seidel's work is unique as a source on African origins. According to Sebro (2010: 67–68), Seidel kept a journal in which he detailed how he was in sustained contact with the enslaved, and he described how he interviewed each member of the congregation separately, presumably with the aid of an interpreter. Nothing indicates that Seidel knew Virgin Islands Dutch Creole, not to mention any African languages, so he would have required the assistance from someone who did. At any rate, Seidel's approach yielded good results. While it would have been possible for the informants to supply only vague information on their ethnic identities, such as "Guinea," the large majority of them chose to self-identify using more salient, lower-level descriptive categories, such as "Allada" or "Watje," which can be linked to modern-day African groups and, by extension, languages.

The information in Seidel's list as tabulated in Sebro (2010: 125–126) is summarized in Table 6. All entries are presented as originally recorded by Seidel. Sebro also includes a number of variant spellings in parentheses, which I have not included. Note that the 36 people who did not state an identifiable group are omitted from the table. Of those whose ethnic self-designations are obscure or lacking altogether, two indicated to have been born during the crossing of the Middle Passage; one simply self-identified as being from "Guinea"; and 33 apparently did not provide an answer to the question, or the answer was not recorded by Seidel. In Sebro's tabulation, a number of individuals can be seen to have used ethnonyms relating to their parents rather than to themselves (e.g., as *geb. St. Thomas Eltern Aja* 'born on St. Thomas to Aja parents'), indicating that they themselves were locally born but continued to identify with their African heritage (*vis-à-vis* the 125 people who self-identified as being "Criol," i.e., locally-born Creoles). Drawing on

secondary sources (e.g., Thornton 1998; Hall 2005; Law 2005), as well as considering information from Oldendorp (2000, 2002) and other primary sources, Sebrot tentatively identified some of the languages/language groups presumed to correlate with the attested ethnonyms. Others are added by me in Table 6, informed primarily by the analysis of Oldendorp’s data (cf. §4). Thirteen designations in the list have not been traced to any specific group.

Table 6: Seidel’s list (1753) summarized, arranged alphabetically

“Nation”	<i>N</i>	%	Correlate language/language group
Aja	13	3.1	Gbe, Kwa
Ajonga	1	<1	–
Akrum	1	<1	–
Allada	4	<1	Gbe, Kwa
Amina	34	8.1	Akan, Kwa
Amombamba	1	<1	Mbamba dialect of Kimbundu, Bantu
Angkrang	1	<1	Gā-Dangme, Kwa
Bundu	1	<1	– (Mandinka, Mande? Kimbundu, Bantu?)
Chamba	13	3.1	Gurma, Gur
Comba	1	<1	– (Kikongo, Bantu, Benue-Congo? Gur? Kru?)
Criol	125	29.9	Virgin Islands Dutch Creole (and presumably other languages)
Esina	1	<1	–
Fanti	2	<1	Akan, Kwa
Fon	7	1.7	Gbe, Kwa
Gango	1	<1	– (Kikongo, Bantu, Benue-Congo? Kru?)
Ibo	8	1.9	Igbo, Benue-Congo
Kamba	1	<1	Kikongo, Bantu, Benue-Congo
Karabari	5	1.2	Igbo, Benue-Congo
Kongo	19	4.5	Kikongo, Bantu, Benue-Congo
Kpesi	2	<1	Gbe, Kwa
Lique	1	<1	Likpe [lip], Kwa
Loango	28	6.7	Kikongo, Bantu, Benue-Congo
Lunda	1	<1	Lunda, Bantu, Benue-Congo
Mandinga	12	2.9	Mandinka, Mande
Mandongo	5	1.2	Kimbundu(?), Bantu, Benue-Congo
Nago	5	1.2	Yoruba, Benue-Congo
Ongokalla	1	<1	–
Ouidah	12	2.9	Gbe, Kwa
Popo	15	3.6	Gbe, Kwa
Poshee	1	<1	–
Rentha	1	<1	–
Rüba	1	<1	–

2 African ethnolinguistic diversity in the colonial Caribbean

“Nation”	N	%	Correlate language/language group
Soko	3	<1	Bambara, Mande
Sundi	1	<1	Kikongo, Bantu, Benue-Congo
Tem	6	1.4	Tem, Gur
Tori	4	1	Gbe, Kwa
Vuningah	1	<1	–
Watje	41	9.8	Gbe, Kwa
Wenwig	1	<1	–
Wungsoko	1	<1	–

Figures 3 and 4 summarize the distributions in terms of the ethnolinguistic correlate groups and the regions of provenance of these, respectively. The two charts highlight the fact that the locally-born Creoles ($N = 125$ out of 418, 29.9%) comprise the largest group in Seidel’s list, which can be seen to reflect a difference between St. Thomas and St. Croix in terms of the proportion of Creole to African enslaved people at this point in time. This is a consequence of the fact that St. Thomas had been colonized more than half a century before St. Croix, and that the population on St. Thomas had better conditions for reproduction. Moreover, the actual share of local-borns was in fact higher than reflected in these figures, as some opted to identify with their African heritage, as discussed above.

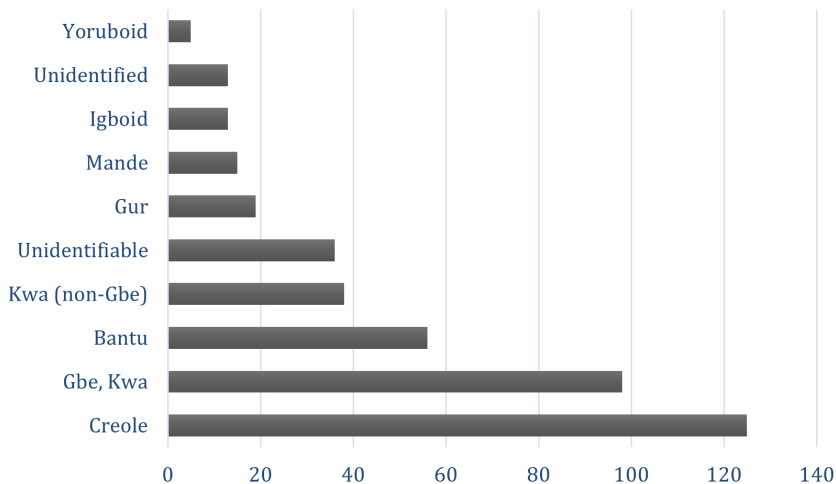


Figure 3: The distribution of language groups summarized

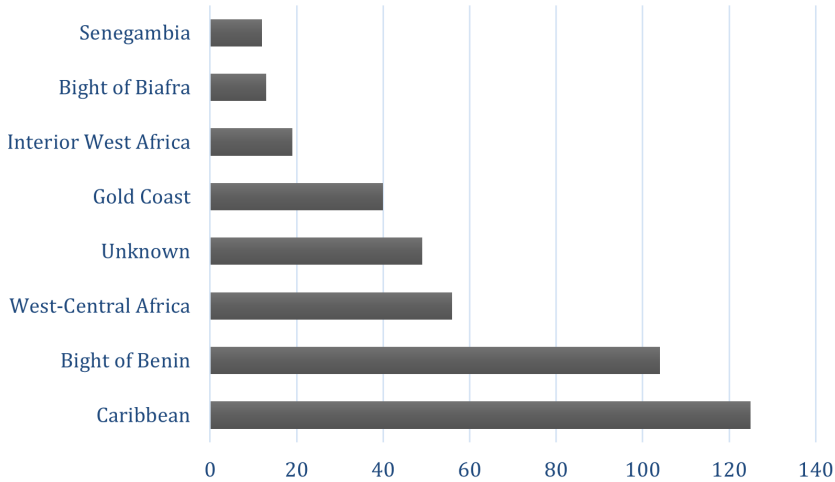


Figure 4: The geographic distribution of the groups summarized

Turning to the African groups, we can note that all of the unidentified “groups” consist of just one individual. This means that they are marginal, only thirteen individuals out of 418 people (i.e., 3.1%).

104 individuals (24.9%) were seemingly related to the Bight of Benin, and 98 of these people can be presumed to have been (heritage) speakers of Gbe lects. Thus, based on its representation in Seidel’s list, it would appear that Sabino (e.g., 1988, 2012) is correct in considering Gbe a likely main substrate language of Virgin Islands Dutch Creole.

The second-largest African group was that of West-Central Africa ($N = 56$, 13.4%), which undoubtedly would have principally included speakers of Bantu languages. A roughly comparable proportion of at least eight (more than 10%) of Oldendorp’s informants were Bantu speakers. This point is noteworthy in light of the documented absence of enslaved people originating from Bantu-speaking regions via the transatlantic slave trade between 1725 and the 1790s. However, it has to be kept in mind that the enslaved people who came to the Danish West Indies from within the Caribbean, e.g., from Curaçao or other transshipment ports, could well have been Bantu speakers. A supposition that Bantu languages played a prominent role in the early language history of Danish St. Thomas can be backed up by means of linguistic evidence. Parkvall (2019) investigated the etymology of 44 African-derived words in Virgin Islands Dutch Creole as recorded by de Josselin de Jong (1926). He found that 35% of the lexical Africanisms under

consideration could be traced to a Bantu source, which must be viewed as a high percentage. By comparison, 29% had a Kwa (including Gbe) source. These figures are particularly instructive if one supposes that early substrate influences would have been those to manifest themselves most clearly in the creole (e.g., Mufwene 1996). When based on figures from the transatlantic trade, the percentage of Bantu-derived lexical items appears disproportional to the number of early enslaved people on St. Thomas taken from Bantu-speaking areas. Based on those data, as noted by Parkvall (2000: 153), “we would only expect Bantu languages to have had a decisive impact [in] the two decades between 1700 and 1720.” Here, Seidel’s list would appear to offer a more realistic reflection of a continued presence of Bantu speakers on St. Thomas.

Coming in third, if we disregard the more numerous “Unknown” category, is the Gold Coast ($N = 40$, 9.6%), which supplied mostly non-Gbe Kwa speakers. Thus, Seidel’s list does not support the idea of (at least a continued) Akan predominance on St. Thomas. The idea of a strong early Akan presence can, however, be backed by linguistic evidence, in the form of substrate features in Virgin Islands Dutch Creole (Parkvall 2000: 153). This apparent discrepancy between demographic and linguistic data can be viewed as weakening the “heavy Kwa bias” hypothesis (cf. §3.1) – especially the form of it assuming a bias in favor of non-Gbe lects. It could be read as pointing to the Akan influence having occurred elsewhere than on St. Thomas (see, e.g., Goodman 1985, who speculated that the earliest enslaved people on St. Thomas brought a Dutch pidgin with them). Although such a scenario would go against the mainstream view on the genesis of Virgin Islands Dutch Creole (see Sabino 2012: 199), it cannot at present be ruled out with certainty.

Finally, Mande-, Igboid-, Gur-, and Yoruboid-speaking people from Senegambia, the Bight of Biafra, and the West African interior all appear to have made up comparatively small groups on St. Thomas in 1753. Yet, when these smaller groups are considered in addition to the more prominent ones discussed above, the overall picture that obtains, is one of a varied population.

In the end, it would appear that there was a (small) preponderance of people originating in Gbe-speaking areas in West Africa on St. Thomas in 1753, provided we disregard the locally-born subset of the population. However, there was no single dominant group or African language. The Bantu presence was stronger than hitherto assumed in the literature (cf. also Bakker 2016a for a discussion of underestimated numbers of Bantus in the Caribbean). It is especially striking that there were more Bantu than non-Gbe Kwa speakers – which was the group traditionally thought to have predominated in the colony. Two caveats have to be mentioned in connection with this assessment. First, it is unknown to what

degree Seidel's sample was representative of the island's enslaved population as a whole; for instance, they were all converted Christians. Second, it cannot be taken for granted that the 1753 snapshot of the substrate population is indicative of the half century or so that came before it. In the next section, I discuss ethnic information over a 90-year period.

6 St. Croix: Baptismal records from the Moravian mission station Friedensthal

After having been acquired from the French in 1733 (as detailed, e.g., by G. H. Høst [1791] in Highfield & Bøegh 2018), St. Croix was settled from 1734 onward, with settlers arriving from St. Thomas, St. John, and elsewhere in the Leeward Islands. St. Croix soon took on the role of the Danish West Indies' most important sugar-producing island. As a result, new arrivals were increasingly sent there, making St. Croix a demographically volatile society characterized by high immigration in the eighteenth century. The enslaved population on St. Croix not only grew much larger than those on St. Thomas and St. John but also became more diversified in terms of its members' ethnolinguistic origins. This conclusion partly reflects the fact that the transatlantic slave trade expanded to encompass a wider range of West African regions in the later eighteenth century, with the primary area of trade shifting eastward and southward on the continent (cf. §3). Even more tellingly, however, it can be corroborated by data on the origins of the enslaved population collected and stored by the Moravian missionaries in St. Croix.

The Moravians established two mission stations on St. Croix in the eighteenth century. In 1734, they built Friedensthal 'The Valley of Peace' at the western end of Christiansted town. In 1771, they opened a second mission station, Friedenberg 'Hill of Peace,' overlooking the newly established town of Frederiksted. Between 1744 and 1832, the Moravians at Friedensthal, the first of these mission stations, baptized a total of 6,783 Africans and African Caribbeans. In the 1960s, the anthropologist Pauline H. Pope carried out archival research on St. Croix for her dissertation (Pope 1970), in which she tabulated data copied directly from the baptismal records stored at the island's Moravian mission stations.¹³ The records have since been moved to Bethlehem, Pennsylvania. Parallel to Seidel in St. Thomas, the missionaries in St. Croix included as part of the record for

¹³In addition to the baptismal records from Friedensthal, we also have data recorded at Friedenberg. In the present study, I focus on the Friedensthal data, as this material extends further back in time than the material kept at Friedenberg.

2 African ethnolinguistic diversity in the colonial Caribbean

each candidate for baptism the name of their “nation.” However, whereas Seidel’s approach was detailed in his journal, the exact method for obtaining this information on St. Croix is unknown. It may be assumed that most missionaries took down information directly from the candidates in question, but, inevitably, many people were involved in this work, so the information may not have as clear an origin in the people it concerned as in the cases of Seidel’s list and Oldendorp’s survey. On the other hand, despite these uncertainties, the scope and volume of the records are both exceptional.

Pope’s (1970) tabulations show that 757 individuals gave no identifiable “nation” or other meaningful designation for the records, or that this information was not taken down. In contrast, 6,026 (or nearly 90%) of the total number of baptismal candidates did provide such information. About 50 different groups appear in the records. Pope attempted to identify these with geographical locations and modern-day African peoples, basing her analysis largely on secondary sources. According to Holsoe (1994: 35), she was for the most part successful, albeit with some exceptions. In subsequent research, Sebros (2010) and Simonsen (2017) have accordingly revised a number of Pope’s conclusions. Summarizing their results, it has turned out that the records, in a number of instances, actually reflect a greater diversity than Pope had recognized. Accordingly, since it is possible to work directly with the source material as tabulated in Pope (1970), I present my own analysis of the baptismal records below, incorporating the mentioned revisions in the process, and adding some of my own.

The data in Table 7 show a distribution by decade¹⁴ for a subsample of the 15 African “nations” which had the largest representation in the Friedensthal records. Figure 5 visualizes changes in the distribution over time.

Notwithstanding the issue of representativeness, the data point in the direction of a predominance of people identified as belonging to the “Mandingo” (i.e., presumably, speakers of the Mande language Mandinka or a closely related lect), “Ibo” (language group: Igboid, Benue-Congo, of Nigeria), “Watyí” (Gbe, Kwa), “Amina” (Akan, Kwa), “Kongo” (Kikongo, Bantu, Benue-Congo), and “Ashanti” (Akan, Kwa) groups. There were likewise concentrations of people identified as “Kalabari” (Igboid, Benue-Congo), “Mokko” (Efik, Cross River, Benue-Congo), “Bambara” (Bambara, Mande, of Mali), and “Akanda,” as well as more modestly sized groups of “Loango” (Kikongo, Bantu, Benue-Congo), “Sokoto” (or “Sokko”:

¹⁴Note the columns for 1744 through 1759 have been conflated in the table since there was only a single individual recorded in the 1744–1749 column. Likewise, the 1830–1832 column only recorded two individuals, and it was therefore merged with the column for 1820–1829. The data in Table 9 (page 53) are arranged the same way.

Table 7: Distribution by decade of 15 major African “nations” represented in the Friedensthal mission records, 1744–1832 (source: Pope 1970: 70–79)

“Nation”	1744–59	1760s	1770s	1780s	1790s	1800s	1810s	1820–32	Totals
Mandingo	6	56	182	165	179	73	19	11	691
Ibo	10	78	156	112	53	53	23	4	489
Waty	16	36	79	91	122	97	35	10	486
Amina	66	84	88	83	67	40	17	1	446
Kongo	17	37	72	61	33	97	51	21	389
Ashanti	21	42	58	96	70	53	18	0	358
Kalabari	29	56	62	36	15	12	4	0	214
Mokko	3	14	43	51	24	32	17	8	192
Bambara	0	13	38	32	50	18	9	0	160
Akanda	0	31	40	48	9	16	3	0	147
Loango	12	34	28	4	3	1	1	1	84
Sokoto	4	18	13	12	9	2	0	0	58
Kwahu	4	7	18	19	8	2	2	0	60
Popo	16	3	1	3	16	2	2	0	43
Nupe	3	9	3	8	0	4	0	0	27

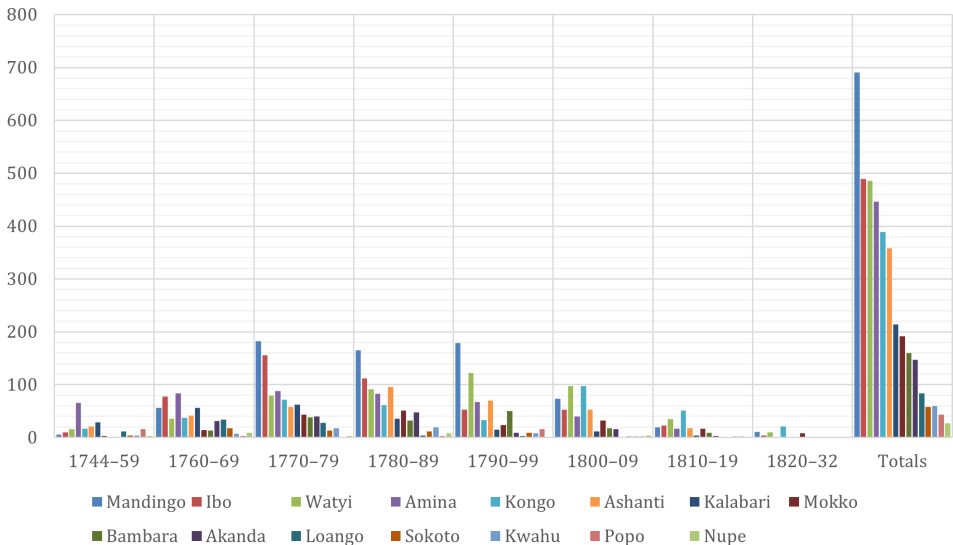


Figure 5: Decade-by-decade and total distribution of 15 major African “nations” present on St. Croix, 1744–1832

2 African ethnolinguistic diversity in the colonial Caribbean

Bambara, Mande), “Kwahu,” “Popo” (Gbe, Kwa), and “Nupe” (presumably Nupoid, Benue-Congo, of Nigeria) people.

The identity of thirteen of the fifteen groups is straightforward. Only the identification of the aforementioned “Akanda” and “Kwahu” groups is problematic. Pope (1970: 25) preferred to link the “Akanda” with Oldendorp’s Kru-speaking “Kanga,” but she also considered the possibility that the “Akanda” were a Yoruboid-speaking group of coastal Nigeria. Alternatively, the group could be traced to an eponymous toponym in Gabon, just as the word may refer to, e.g., Akan or Nupoid groups. In the absence of direct linguistic evidence, I opt not to specify a group. It is likewise not possible to ascertain whether the “Kwahu” group should be counted among the Gbe (i.e., corresponding to Oldendorp’s “Wavu”) or not (cf. Fodor’s 1975 “Wavu II”). In addition, the term might refer to the Twi-speaking Kwahu group of Ghana. Thus, both terms are ambiguous.

A preponderance of “Amina” people ($N = 66$, 32%) can be identified early on (1744–1759), but a shift can be seen from the 1760s onward, in the direction of a more heterogeneous population. Incidentally, the point about early “Amina” (Akan) dominance aligns with another source on the demographic composition of early Danish St. Croix – specifically, a roster of Company-owned enslaved individuals working in Christiansted from 1740 to 1755, sourced from the Danish National Archives (tabulated in Tyson 2011: 45–48). Conversely, the “Amina” are absent from later tallies of a similar kind (e.g., Holsoe 1994: 36).

Table 8 provides a summary of the African language groups that, based on the analysis in §4 and the data presented above, are tentatively identified as being most prominently represented on St. Croix between 1744 and 1832.

Table 8: Tentative identification of modern language groups associated with the 15 major African “nations” on St. Croix, 1744–1832

Language group	“Nation(s)”	<i>N</i>	%
Mande	Mandingo, Bambara, Sokoto	909	24
Akan, Kwa	Amina, Ashanti	804	21
Igboid, Benue-Congo	Ibo, Kalabari	703	18
Gbe, Kwa	Watyí, Popo	529	14
Bantu, Benue-Congo	Kongo, Loango	473	12
–	Akanda, Kwahu	207	5
Cross River, Benue-Congo	Mokko	192	5
Nupoid, Benue-Congo	Nupe	27	1

The subsample considered above, it should be emphasized, offers no insights regarding the Caribbean-born subset of the population, comprising 2,014 (or 33%

of 6,026) people. The African “nations” identified thus far account for 3,844 individuals (64%) out of 6,026 possible. Thus, the distribution seen above cannot be generalized to the full sample. Moreover, it renders invisible the many “nations” that were represented by just a few individuals, such as the “Bulom” ($N = 2$) or the “Pu” ($N = 1$), as well as a number of more notable groups, numerically speaking. Such groups include the “Fulani” ($N = 24$), speakers of Fula lects (North Atlantic, interior Senegambia and West Africa); the “Kissi” ($N = 15$) and “Timne” ($N = 20$), speakers of Mel lects (Sierra Leone and environs); the “Akkran” ($N = 19$), Gã-Dangme speakers of Ghana; and the “Kamba” ($N = 19$) and “Mondonga” ($N = 10$), both presumably speakers of Bantu languages. The remaining African groups all number less than ten individuals, and for the most part less than three. Factoring in these smaller groups only adds further detail to the overall picture of a slave society characterized by considerable ethnolinguistic diversity.

Next, considering the full sample ($N = 6,026$), the data presented in Table 9 show the distribution of the Friedensthal baptismal candidates’ regional origins. Note that I essentially retain Pope’s (1970) geographical categories, albeit with updated terminology, and with certain changes in terms of the specific composition of the groups. For instance, one of Sebro’s (2010: 90) objections to Pope’s analysis is that the “Gold Coast” category was used in a way where it extended further into the West African interior than this label would otherwise rightly be used to denote. It also encompassed predominantly Gur-speaking areas in the interior, thus reflecting a larger diversity than first assessed. Moreover, Pope placed a number of groups that presumably were Yoruboid-speaking in her “Nigerian-Cameroon” (or Bight of Biafra) category, where it would be more accurate to trace these to the Bight of Benin (Pope’s “Slave Coast”). In my analysis, the Gold Coast category has been retained (with the implication that it extends further into the interior than is customary in the literature), but the Yoruboid-speaking groups (comprising 159 individuals) have been counted among the ones originating in the Bight of Benin. Figure 6 offers a visualization of changes in the distribution over time.

Based on these data, the single largest group for the full period covered is the Caribbean-born subset of the population, which includes those born in the Danish West Indies ($N = 1,445$, 24%), as well as individuals taken to St. Croix from elsewhere in the West Indies/Caribbean region ($N = 569$, 12%). Nigeria-Cameroon emerges as the African region most clearly represented in the baptismal records ($N = 985$, 16%), followed closely by the Gold Coast ($N = 904$, 15%) and Senegambia ($N = 884$, 15%). Next in line are the Bight of Benin ($N = 690$, 9%), West-Central Africa ($N = 509$, 8%), and the coastal stretch along Sierra Leone,

2 African ethnolinguistic diversity in the colonial Caribbean

Table 9: Regional origins of the enslaved population on St. Croix, 1744–1832 (source: Pope 1970: 59–69)

Region	1744–59	1760s	1770s	1780s	1790s	1800s	1810s	1820–32	Totals
Danish West Indies	91	48	221	222	283	245	182	119	1,445
Nigeria-Cameroon	49	178	277	220	101	104	44	12	985
Gold Coast	98	143	173	204	150	100	35	1	904
Senegambia	6	71	227	203	238	96	30	13	884
Bight of Benin	23	87	122	140	136	129	40	13	690
West Indies (elsewhere)	25	153	224	111	40	8	5	3	569
West-Central Africa	26	85	106	72	44	100	52	24	509
Sierra Leone, Liberia, Ivory Coast	0	0	2	3	11	17	5	3	41

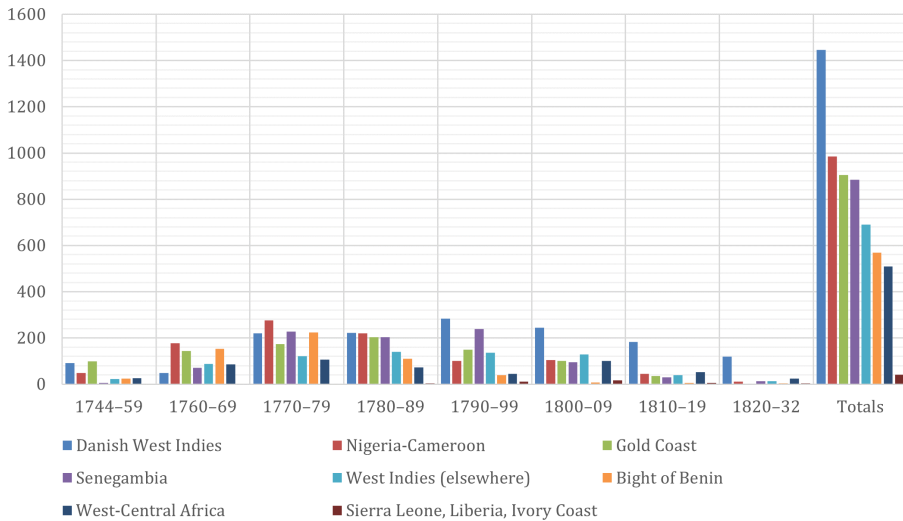


Figure 6: Decade-by-decade and total distribution of the regional origins of Africans and African Caribbeans baptized at Friedensthal, 1744–1832

Liberia, and the Ivory Coast ($N = 41$, 1%). These numbers reflect enormous diversity (even disregarding the diversity within these regions), with no groups getting close to a majority.

If one considers diachronic developments in the distribution, it is noteworthy that the locally-born population has a strong presence mainly from the 1770s onward, and a predominance only from the 1790s. This indicates that the overall population became more and more locally born, but also that this development took place later and at a slower pace than on St. Thomas. Alongside the locally-born population, individuals originating in Africa continued to be baptized in large numbers throughout the eighteenth century, and, to a lesser degree, well into the nineteenth century. This is in line with the conclusion of Simonsen & Olsen (2017: 151) who, writing on overall tendencies in demographic developments in the Danish West Indies, state the following: “For St. Croix, thus, there was a marked Africanization of the island, understood in the sense that the meeting of Africans with different linguistic, ethnic, and religious affiliations came to permeate life on the island” (my translation).

7 Summary and concluding remarks

This study has sought to reconstruct a profile of the African ethnolinguistic groups represented in the eighteenth-century Danish West Indies. I examined quantitative data on the transatlantic and intra-Caribbean slave trade to the colony, as well as qualitative and quantitative documentation collected by Moravian missionaries active in the islands. I compared the linguistic data recorded by Oldendorp with modern data on hundreds of languages spoken in West and West-Central Africa, and linked the “nations” identified on this basis with independent data on ethnic origins of the enslaved people on the islands. Those ethnic data were represented in the list of Christianized enslaved individuals from St. Thomas and in baptismal records from St. Croix. I found that it was altogether possible to set out the sources of enslaved people and their languages or language groups represented in the colony.

The different pieces of the puzzle consistently point in the direction of a greater ethnolinguistic diversity than had been assumed to be the case for the Danish West Indies in previous contact linguistic research. The view that a single ethnolinguistic group (Kwa speakers) predominated in the colony in the eighteenth century cannot be upheld, at least not categorically. Enslaved Africans were taken from multiple localities, including from far into the interior. Many of the languages identified are not coastal, hence shipment ports are potentially

misleading. In addition to Kwa languages, it appears that Bantu and Nigerian languages – and others yet – had a solid representation in the enslaved population. This indicates, among other things, that the sociohistorical context in which creole languages developed within Virgin Islands society during the colonial period was characterized by significant linguistic diversity among the enslaved. This indicates that the set of potential substrate languages for Virgin Islands Dutch Creole (and, in a more limited, adstratal way, the English-based varieties of the islands) should be expanded. With respect to the question of ethnolinguistic diversity, it is worth considering whether the Danish colony was in any way exceptional – besides the fact that a substantial amount of data has survived. If indeed the overall tendencies discerned for the Danish West Indies apply to other Caribbean islands that share a similar history, then it follows that a number of other territories can look to the Virgin Islands for a more realistic view of the extensiveness of their ethnolinguistic diversity throughout the colonial period.

An important caveat to the conclusions presented here is that the enslaved Africans were generally multilingual, speaking one or more regional lingua francas as well. This study has not addressed this issue, mainly because the data considered offer few insights into it. Another limitation concerns potential underrepresentation of certain groups in the sources considered, which cannot be ruled out. For instance, some African groups may have resisted Christianization, and those would be underrepresented.

Acknowledgements

I wish to thank Peter Bakker and two anonymous reviewers for their valuable feedback on an earlier version of this study. I am also grateful to the audience at my presentation at the 2017 Society for Pidgin and Creole Linguistics' Summer Meeting in Tampere, Finland, for their comments on my initial work on this topic. When finalizing this study, I had the benefit of a grant from the Carlsberg Foundation (Grant CF23-1162).

References

Ahlbäck, Anders. 2016. The overly candid missionary historian: C.G.A. Oldendorp's theological ambivalence over slavery in the Danish West Indies. In Holger Weiss (ed.), *Ports of globalisation, places of creolisation: Nordic possessions in the Atlantic world during the era of the slave trade*, 191–217. Leiden: Brill. DOI: 10.1163/9789004302792_008.

- Arends, Jacques. 2008. A demographic perspective on creole formation. In Silvia Kouwenberg & John Victor Singler (eds.), *The handbook of pidgin and creole studies*, 309–331. Oxford: Wiley-Blackwell.
- Arends, Jacques & Pieter Muysken. 1992. Demografische Modellen in de Creolistiek. *Gramma/TTT, tijdschrift voor taalkunde* 1(1). 41–55.
- Asmussen, Jesper Bering. 1983. *Slavedemografi. St. Croix' landdistrikter, 1803–1848* [*Slave demography. St. Croix's land districts, 1803–1848*]. Aarhus Universitet. (MA Thesis).
- Auerbach, Joh. Christoph. 1774. [*Letter from Niesky in St. Thomas, 10 March 1774 to Friedrich Neisser, Barby, Germany.*] Manuscript.
- Bakker, Peter. 2016a. Creoles as hybrid languages. *Word* 62(4). 228–243. DOI: 10.1080/00437956.2016.1248655.
- Bakker, Peter. 2016b. Review of language contact in the Danish West Indies: Giving Jack his jacket. By Robin Sabino. *Journal of Pidgin and Creole Languages* 31(1). 223–228. DOI: 10.1075/jpcl.31.1.11bak.
- Ballester, Maria Teresa Galarza. 2016. A socio-historical account of the formation of the creole language of Antigua. *Journal of Pidgin and Creole Languages* 31(2). 288–315. DOI: 10.1075/jpcl.31.2.03gal.
- Bartens, Angela & Philip Baker. 2012. *Black through white: African words and calques in creoles and transplanted European languages*. London: Battlebridge.
- Bello, Mariama. 2013. *Assessing language vitality and language endangerment of Lete (Larteh)*. University of Ghana. (MA thesis).
- Birmingham, David. 1966. *Trade and conflict in Angola: The Mbundu and their neighbours under the influence of the Portuguese 1483–1790*. Oxford: Clarendon Press.
- Bøegh, Kristoffer Friis & Peter Bakker. 2021. Words from Dutch Creole in Virgin Islands Creole English. *Trefwoord (2021)*. 1–38. <https://neerlandistiek.nl/2021/09/nieuw-in-trefwoord-3/>.
- Bøegh, Kristoffer Friis, Peter Bakker, Cefas van Rossem & Rasmus Christensen. 2022. Seven newly discovered 18th and 19th century Virgin Islands Dutch Creole texts. In Nicholas Faraclas, Ronald Severing, Sally Delgado, Elisabeth Echteld & Wim Rutgers (eds.), *Caribbean convivialities and Caribbean sciences: Inclusive approaches to the study of the languages, literatures and cultures of the Dutch Caribbean and beyond*, 93–116. Willemstad: University of Curaçao.
- Bøegh, Kristoffer Friis, Aymeric Daval-Markussen & Peter Bakker. 2016. A phylogenetic analysis of stable structural features in West African languages. *Studies in African Linguistics* 45(1). 62–94. DOI: 10.32473/sal.v45i1.107255.
- Bureau of Ghana Languages. 1986. *Language guide: Fante*. Accra: Bureau of Ghana Languages.

2 African ethnolinguistic diversity in the colonial Caribbean

- da Costa Peixoto, António. 1944. *Obra nova de língua geral de Mina*. Luís Silveira (ed.). Lisboa: Agência Geral das Colónias.
- Dalby, David. 1964. Provisional identification of languages in the Polyglotta Africana. *Sierra Leone Language Review* 3. 83–90.
- de Josselin de Jong, J.P.B. 1926. *Het huidige Negerhollandsch (Texten en woordenlijst)*. Amsterdam: Koninklijke Akademie van Wetenschappen te Amsterdam.
- Dryer, Matthew S. 2005. Genealogical language list. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 584–644. Oxford: Oxford University Press.
- Dryer, Matthew S. & Martin Haspelmath. 2013. *The world atlas of language structures online*. Leipzig. <http://wals.info> (30 September, 2019).
- Eberhard, David M., Gary F. Simons & Charles D. Fenning. 2019. *Ethnologue: Languages of the world. Twenty-second edition*. Dallas, Texas: SIL International.
- Eltis, David, Stephen D. Behrendt, David Richardson & Herbert S. Klein. 1999. *The transatlantic slave trade. A database on CD-ROM*. Cambridge: Cambridge University Press.
- Feldbæk, Ole & Ole Justesen. 1980. *Kolonierne i Asien og Afrika [The colonies in Asia and Africa]*. Copenhagen: Politiken.
- Fodor, István. 1975. *Pallas und andere afrikanische Vokabularien vor dem 19. Jahrhundert: Ein Beitrag zur Forschungsgeschichte*. Hamburg: Helmut Buske.
- Gøbel, Erik. 2016. *The Danish slave trade and its abolition*. Leiden/Boston: Brill.
- Gøbel, Erik & Louise Sebro. 2017. Danmark-Norge i Vestindien 1672–1720 og Forandringens år 1720–1740 [Denmark-Norway in the West Indies 1672–1720 and Years of change 1720–1740]. In Erik Gøbel & Louise Sebro (eds.), *Vestindien. St. Croix, St. Thomas og St. Jan [The West Indies. St. Croix, St. Thomas, and St. John]*, 48–129. Copenhagen: Gads Forlag.
- Goodman, Morris F. 1985. The origin of Virgin Islands Dutch Creole. *Amsterdam Creole Studies* 8. 67–106.
- Green-Pedersen, Svend E. 1971. The scope and structure of the Danish Negro slave trade. *Scandinavian Economic History Review* 19(2). 149–197. DOI: 10.1080/03585522.1971.10407697.
- Green-Pedersen, Svend E. 1975. The history of the Danish Negro slave trade 1733–1807: An interim survey relating in particular to its volume, structure, profitability and abolition. *Revue Francaise d’Histoire d’Outre Mer* 62. 196–220.
- Green-Pedersen, Svend E. 1981. Slave demography in the Danish West Indies and the abolition of the Danish slave trade. In David Eltis & James Walvin (eds.), *The abolition of the Atlantic slave trade. Origins and effects in Europe, Africa, and the Americas*, 231–257. Madison, WI: University of Wisconsin Press.

- Güldemann, Tom. 2018. *The languages and linguistics of Africa*. Berlin/Boston: De Gruyter Mouton. DOI: 10.1515/9783110421668.
- Guthrie, Malcolm. 1948. *The classification of the Bantu languages*. London: Oxford University Press for the International African Institute.
- Hair, Paul E. H. 1967. Ethnolinguistic continuity on the Guinea Coast. *The Journal of African History* 8(2). 247–268.
- Hair, Paul E. H. 1968. An ethnolinguistic inventory of the lower Guinea coast before 1700. Pt. 1. *African Language Review* 7. 113–121.
- Hall, Gwendolyn M. 2005. *Slavery and African ethnicities in the Americas: Restoring the links*. Chapel Hill: The University of North Carolina Press.
- Hall, Neville T. 1992. *Slave Society in the Danish West Indies: St. Thomas, St. John and St. Croix*. Mona, Jamaica: University of the West Indies Press.
- Herskovits, Melville L. 1958. *The myth of the Negro past*. Boston: Beacon Press.
- Highfield, Arnold R. 1979. *The French dialect of St. Thomas, United States Virgin Islands: A descriptive grammar with texts and glossary*. Ann Arbor: Karoma Press.
- Highfield, Arnold R. 1993. Toward a language history of the Danish West Indies and the United States Virgin Islands. In Svend E. Holsoe & John H. McCollum (eds.), *The Danish presence and legacy in the Virgin Islands*, 122–139. Frederiksted, St. Croix: St. Croix Landmarks Society & the V.I. Humanities Council.
- Highfield, Arnold R. 1994. Patterns of accommodation and resistance: The Moravian witness to slavery in the Danish West Indies. *Journal of Caribbean History* 28(2). 138–164.
- Highfield, Arnold R. & Vladimir Barac. 1987. *C. G. A. Oldendorp's history of the Mission of the Evangelical Brethren on the Caribbean islands of St. Thomas, St. Croix, and St. John*. Ann Arbor: Karoma Publishers.
- Highfield, Arnold R. & Kristoffer Friis Bøegh. 2018. *Georg Hjersing Høst's account of the island of St. Thomas and its governors. Recorded there on the island from 1769 until 1776*. Christiansted, St. Croix, USVI: Antilles Press.
- Higman, B. W. 2011. *A concise history of the Caribbean*. Cambridge: Cambridge University Press.
- Holsoe, Svend E. 1994. The origin, transport, introduction and distribution of Africans on St. Croix: An overview. In George F. Tyson & Arnold R. Highfield (eds.), *The Danish West Indian slave trade: Virgin Islands perspectives*, 33–46. St. Croix: Virgin Islands Humanities Council.
- Hopkins, Daniel P. 1987. *The Danish cadastral survey of St. Croix, 1733–1754*. Louisiana State University. (Doctoral dissertation).
- Hornby, Ove. 1980. *Kolonierne i Vestindien [The colonies in the West Indies]*. Copenhagen: Politiken.

2 African ethnolinguistic diversity in the colonial Caribbean

- Hvid, Mirjam Louise. 2016. Indentured servitude and convict labour in the Danish-Norwegian West Indies, 1671–1755. *Scandinavian Journal of History* 41(4-5). 541–564. DOI: 10.1080/03468755.2016.1210890.
- Innes, Gordon. 1967. *A Grebo-English dictionary*. Cambridge: Cambridge University Press.
- Jensen, Niklas Thode & Gunvor Simonsen. 2016. Introduction: The historiography of slavery in the Danish-Norwegian West Indies, c. 1950–2016. *Scandinavian Journal of History* 41(4-5). 475–494. DOI: 10.1080/03468755.2016.1210880.
- Jones, Adam. 1991. *Zur Quellenproblematik der Geschichte Westafrikas*. Stuttgart: Franz Steiner Verlag.
- Jones, Adam. 2010. Oldendorps Beitrag zur Afrikaforschung. In Gudrun Meier (ed.), *Christian Georg Andreas Oldendorp, Historie der Caraibischen Inseln Sanct Thomas, Sanct Crux und Sanct Jan: Kommentarband*, 181–190. Herrnhut: Herrnhuter Verlag.
- Kelley, Sean M. & Paul E. Lovejoy. 2023. Oldendorp’s ‘Amina’: Ethnonyms, history, and identity in the African diaspora. *Journal of Global Slavery* 8. 303–330. DOI: 10.1163/2405836X-00802015.
- Koelle, Sigismund W. 1854. *Polyglotta Africana*. London: London Church Missionary House.
- Kouwenberg, Silvia. 2008. The problem of multiple substrates: The case of Jamaican Creole. In Susanne Michaelis (ed.), *Roots of creole structures. Weighing the contribution of substrates and superstrates*, 1–27. Amsterdam/Philadelphia: John Benjamins.
- Kouwenberg, Silvia & John Victor Singler. 2018. Creolization in context: Historical and typological perspectives. *Annual Review of Linguistics* 4. 213–232. DOI: 10.1146/annurev-linguistics-011415-040544.
- Larsen, Jens. 1950. *Virgin Islands story*. Philadelphia, PA: Fortress Press.
- Law, Robin. 2005. Ethnicities of enslaved Africans in the diaspora: On the meanings of “Mina” (again). *History in Africa* 32. 247–267. DOI: 10.1353/hia.2005.0014.
- Lawaetz, H. 1902. *Brødremenighedens Mission i Dansk-Vestindien 1769–1848: Bidrag til en Charakteristik af Brødremenigheden og dens Gerning og af den Farvede Races Stilling til Christendommen [The Brethren’s congregation in the Danish West Indies 1769–1848: Contribution to a characterization of the Brethren’s congregation and its acts and of the colored race’s attitude to Christianity]*. Copenhagen: Otto B. Wroblewski.
- Lefebvre, Claire & Anne-Marie Brousseau. 2002. *A grammar of Fongbe*. Berlin: Mouton de Gruyter.

- Magens, J.M. 1770. *Grammatica over det Creolske Sprog, som bruges paa de trende Danske Eilande, St. Croix, St. Thomas og St. Jans i America* [Grammar of the Creole language used on the three Danish islands of St. Croix, St. Thomas, and St. John in America]. Copenhagen: Gerhard Giese Salikath.
- Meier, Gudrun, Peter Stein, Stephan Palmié & Horst Ulbricht. 2010. *Christian Georg Andreas Oldendorp, Historie der caribischen Inseln Sanct Thomas, Sanct Crux und Sanct Jan. Kommentarband*. Herrnhut: Herrnhuter Verlag.
- Migeod, Frederick W. H. 1913. *The languages of West Africa. Part 2*. London: Kegan Paul, Trench, Trubner & Co.
- Mufwene, Salikoko S. 1996. The founder principle in creole genesis. *Diachronica* 13(1). 83–134. DOI: 10.1075/dia.13.1.05muf.
- Muysken, Pieter, Norval S. H. Smith & Robert Borges. 2015. *Surviving the Middle Passage: The West Africa-Surinam Sprachbund*. Berlin/Boston: De Gruyter Mouton. DOI: 10.1515/9783110343977.
- Nielsen, Herluf. 1981. *Dansk Vestindien for 250 år siden. J. L. Carstens: En Almindelig Beskrivelse om alle de Danske, Americanske eller West-Indiske Ey-Lande* [The Danish West Indies 250 years ago. J. L. Carstens: An ordinary description concerning all the Danish, American or West-Indian islands]. Herluf Nielsen (ed.). Copenhagen: Dansk Vestindisk Forlag.
- Nørregaard, Georg. 1966. *Danish settlements in West Africa 1658–1850*. Boston: Boston University Press.
- Oldendorp, Christian Georg Andreas. 1777. *Geschichte der Mission der evangelischen Brüder auf den caraibischen Inseln S. Thomas, S. Croix und S. Jan*. J. J. Bossart (ed.). Barby: Laux.
- Oldendorp, Christian Georg Andreas. 1996. *Criolisches Wörterbuch. (Lexicographica: Series maior, 69)*. Peter Stein & Hein van der Voort (eds.). Tübingen: Niemeyer.
- Oldendorp, Christian Georg Andreas. 2000. *Historie der caribischen Inseln Sanct Thomas, Sanct Crux und Sanct Jan, insbesondere der dasigen Neger und der Mission der evangelischen Brüder under derselben. Erster Teil*. Hartmut Beck, Gudrun Meier, Stephan Palmié, Peter Stein & Horst Ulbricht (eds.) (Abhandlungen und Berichte des Staatlichen Museums für Völkerkunde Dresden 51). Berlin: VWB Verlag für Wissenschaft und Bildung.
- Oldendorp, Christian Georg Andreas. 2002. *Historie der caribischen Inseln Sanct Thomas, Sanct Crux und Sanct Jan, insbesondere der dasigen Neger und der Mission der evangelischen Brüder under derselben. Zweiter Teil*. Hartmut Beck, Gudrun Meier, Stephan Palmié, Peter Stein & Horst Ulbricht (eds.), 3 vols. (Abhandlungen und Berichte des Staatlichen Museums für Völkerkunde Dresden 51). Berlin: VWB Verlag für Wissenschaft und Bildung.

2 African ethnolinguistic diversity in the colonial Caribbean

- Olsen, Poul Erik. 1988. *Toldvæsenet i Dansk Vestindien 1672–1917* [The customs service in the Danish West Indies 1672–1917]. Copenhagen: Toldhistorisk Selskab.
- Olsen, Poul Erik. 2017. *Vestindien: St. Croix, St. Thomas og St. Jan* [The West Indies: St. Croix, St. Thomas, and St. John]. Copenhagen: Gads Forlag.
- Parkvall, Mikael. 2000. *Out of Africa: African influences in Atlantic creoles*. London: Battlebridge Publications.
- Parkvall, Mikael. 2019. *Afrolex: Africal lexical influences in creoles and transplanted European languages*. Manuscript/database.
- Peace Corps The Gambia. 1995. *Mandinka – English dictionary*. Banjul: The Gambia.
- Pontoppidan, Erik. 1881. Einige Notizen über die Kreolsprache der Dänisch-Westindischen Inseln. *Zeitschrift für Ethnologie* 13. 130–138.
- Pope, Pauline Holman. 1970. *Cruzan slavery: An ethnohistorical study of differential responses to slavery in the Danish West Indies*. Davis: University of California, Davis. (Doctoral dissertation).
- Prescod, Paula & Adrian Fraser. 2015. Sociohistorical and linguistic account of St Vincent and the Grenadines. In Paula Prescod (ed.), *Language issues in Saint Vincent and the Grenadines*, 1–44. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Roberts, Justin, Philip D. Morgan & Rasmus Christensen. 2024. The paradox of abolition: Sugar production and slave demography in Danish St. Croix, 1792–1804. *Journal of Interdisciplinary History* 54(4). 453–476. DOI: 10.1162/jinh_a_02006.
- Sabino, Robin. 1988. The copula in vernacular Negerhollands. *Journal of Pidgin and Creole Languages* 3(2). 199–212. DOI: 10.1075/jpcl.3.2.04sab.
- Sabino, Robin. 2012. *Giving Jack his jacket: Language contact in the Danish West Indies*. Leiden: Brill Publishing.
- Sebro, Louise. 2010. *Mellem afrikaner og kreol: Etnisk identitet og social navigation i Dansk Vestindien 1730–1770* [Between African and Creole: Ethnic identity and social navigation in the Danish West Indies, 1730–1770]. Lund: Lunds Universitet.
- Sebro, Louise. 2016. *Freedom, autonomy, and independence: Exceptional African Caribbean life experiences in St. Thomas, the Danish West Indies, in the middle of the 18th century*. Holger Weiss (ed.). Leiden: Brill. 218–244.
- Seidel, Nathanael. 1753. *Zeugniße von den Communicanten unser Neger-Gemeine in S. Thomas verfertigt von Bruder Nathanael in Gemeinschaft der dortigen Arbeiter bey seiner Visitation in S. Thomas im Monat Junio und Julio 1753*. Manuscript.
- Sensbach, Jon F. 2005. *Rebecca's revival. Creating Black Christianity in the Atlantic world*. Cambridge: Harvard University Press.

- Simonsen, Gunvor. 2017. *Slave stories: Law, representation, and gender in the Danish West Indies*. Aarhus: Aarhus University Press.
- Simonsen, Gunvor & Poul Erik Olsen. 2017. Slavesamfundet konsolideres 1740–1802 [The slave society is consolidated 1740–1802]. In Poul Olsen (ed.), *Vestindien: St. Croix, St. Thomas og St. Jan [The West Indies: St. Croix, St. Thomas, and St. John]*, 130–208. Copenhagen: Gads Forlag.
- Stahl, Ann B. 2001. *Making history in Banda: Anthropological visions of Africa's past*. Cambridge: Cambridge University Press.
- Stein, Peter. 2014. Dr. med. Erik Pontoppidan und das Negerhollands. In Philip Krämer (ed.), *Ausgewählte Arbeiten der Kreolistik des XIX. Jahrhunderts. Selected works from XIXth century creolistics. Emilio Teza, Thomas Russell, Erik Pontoppidan, Adolpho Coelho*, 109–121. Hamburg: Buske.
- Stolz, Thomas & Peter Stein. 1986. Language and history in the former Danish Antilles: Non-linguistic evidence for a diachronic description of the Negro Dutch language. *Amsterdam Creole Studies* 9. 103–122.
- Thornton, John K. 1998. *Africa and Africans in the making of the Atlantic world*. 2nd edn. Cambridge: Cambridge University Press.
- Tyson, George. 2011. Getting it straight: The contributions of Africans to the establishment of Christiansted. In David Hayes (ed.), *Christiansted at 275: Celebrating the 275th anniversary of the founding of Christiansted, St. Croix*, 29–50. St. Croix, US VI: The Society of Virgin Islands Historians.
- van Rossem, Cefas. 2017. *The Virgin Islands Dutch Creole textual heritage: Philological perspectives on authenticity and audience design*. Utrecht: LOT.
- van Rossem, Cefas & Hein van der Voort. 1996. *Die Creol Taal: 250 years of Negerhollands texts*. Amsterdam: Amsterdam University Press.
- van Sluijs, Robert. 2017. *Variation and change in Virgin Islands Dutch Creole: Tense, modality and aspect*. Nijmegen: Radboud Universiteit. (Doctoral dissertation).
- Velupillai, Viveka. 2015. *Pidgins, creoles and mixed languages: An introduction*. Amsterdam/New York: Benjamins.
- Vos, Jelmer. 2010. The slave trade from the Windward Coast: The case of the Dutch, 1740–1805. *African Economic History* 38. 29–51.
- Weindl, Andrea. 2008. The slave trade of Northern Germany. In David Eltis & David Richardson (eds.), *Extending the frontiers: Essays on the new Transatlantic Slave Trade Database*, 250–271. New Haven/London: Yale University Press.
- Westergaard, Waldemar. 1917. *The Danish West Indies under Company rule*. New York: Macmillan Company.
- Westermann, Diedrich H. & M. A. Bryan. 1952. *Languages of West Africa*. Oxford: Oxford University Press.

2 African ethnolinguistic diversity in the colonial Caribbean

Williamson, Kay & Roger Blench. 2000. Niger-Congo. In Bernd Heine & Derek Nurse (eds.), *African languages: An introduction*, 11–42. Cambridge: Cambridge University Press.

Chapter 3

The expression of motion events in Haitian Creole

Carolin Ulmer

Freie Universität Berlin

This paper investigates the expression of motion events in Haitian Creole. A bipartite typology has been proposed by Talmy (1991), sorting languages into verb-framed and satellite-framed languages, depending on where they express the Path component of motion events. Slobin (2004) expanded the typology by a third type, equipollently-framed languages, to include verb-serializing languages which can express the Path as well as the Manner component in a serial verb construction. Creole languages have so far received little to no attention in regard to this typology. Creole languages are especially interesting because they were formed in a situation of language contact. The investigation of their morphosyntactic features can shed light on the question of which features of the languages involved are passed on and which are not. This can in turn offer clues for the study of the markedness of these features. The languages which were relevant to the formation of Haitian Creole, French and Kwa languages, present different patterns here. In French, verb-framed patterns are predominantly used, but in some cases Manner verbs constitute the main verb of the sentence (Pourcel & Kopecka 2005). In contrast, Kwa languages can use verb serializations to encode motion events (Ameka & Essegbey 2013; Lambert-Brétière 2009), a pattern not found in French. In this paper, I describe a small study conducted in Berlin, Germany, in 2017, investigating the expression of Motion events by four native speakers of Haitian Creole. They narrated a picture story and described drawings depicting different combinations of Manner and Path components. A wide range of different morphosyntactic structures encoding motion events was elicited. Verb-framed patterns were frequently used, as well as different Manner-Path verb serializations. Only a few satellite-framed constructions were elicited, but using different Manner verbs and Path-PPs. Further research will need to test the acceptability of different Manner and Path elements in the particular structures.



1 Introduction

The expression of motion events in different languages has been of great interest to many linguists since Talmy (1991) proposed his typology of them, sorting languages into two types depending on whether they typically express the Path of motion in the main verb or a so-called satellite, some other element that is closely associated with the main verb. Romance languages are often cited as typical members of the first group, called verb-framed languages, whereas Germanic languages represent the second group, named satellite-framed languages. Later, Slobin (2004) proposed a third type that he calls equipollently-framed languages to describe languages expressing both the Path and the Manner component in a verb serialization, a structure that is found e.g. in Mandarin Chinese. Even though many languages have been investigated with regard to the morphosyntactic structures used to encode the different components of motion events, there is still no research on this question for Romance-based Creole languages. As Creole languages were formed in a situation of language contact, their investigation can show which features of the languages involved were passed on. The present paper looks at the morphosyntactic expression of Motion events in Haitian Creole. A small pilot study was conducted with four native speakers of Haitian Creole in Berlin, Germany. After a short sociolinguistic interview to determine their language ideologies and habits of language use (which were deemed necessary as the speakers all lived away from their home country and in a multilingual environment), the speakers completed two different tasks. First, they narrated a picture story about a little bird flying out of its cage and house to explore the outside world. After that, they provided descriptions for single pictures which were assembled in order to control for several combinations of Manner and Path components of motion events. The results show that Haitian Creole possesses a rich inventory of morphosyntactic structures to express motion events. A preference exists for the use of verb-framed constructions, but Manner-Path verb serializations were also used frequently. Satellite-framed constructions were rare, but do not seem to be totally ungrammatical.

The structure of the paper is as follows. Section 2 first gives a definition of motion events and then describes the three types of motion event encoding mentioned above. Then, the concept of Manner salience, which describes the frequency with which Manner elements are used in different languages, is introduced. Sections 2.3 and 2.4 provide an insight into the expression of motion events in French and some African languages which were relevant to the formation of Haitian Creole. After that, a first light is shed on motion event encoding in Haitian Creole from previous works on the Haitian Creole language. Section 3

then outlines the study conducted by the author and explains how the data was classified. Section 4 gives a broad overview of the different structures that were elicited, which are then discussed in Section 5.

2 Motion events

In this study, a motion event is understood as defined by Talmy (1991: 60–61):

The basic motion event consists of one object (the “Figure”) moving or located with respect to another object (the reference object or “Ground”). It is analyzed as having four components: besides “Figure” and “Ground”, there are “Path” and “Motion”. The “Path” [...] is the course followed or site occupied by the Figure object with respect to the Ground object. [...] In addition to these internal components a Motion event can have a “Manner” or a “Cause”, which we analyze as constituting a distinct external event.

In sentence (1), the components of the motion event are distributed as follows:

- (1) Tom is running down the stairs.
Manner Path

Tom represents the Figure, *running* the Manner, *down* the Path, and *the stairs* the Ground of the motion event described.

Different lexicalization patterns are found in the languages of the world concerning the motion event component expressed in the verb. A first type encodes Motion and Manner in the verb, which is typical e.g. in English, as seen in sentence (1). A second type, which is typically found in Romance languages like French or Spanish, encodes Motion and Path in verbs, such as *descendre/bajar* (‘go.down’) (Talmy 1985: 62–68). These differences lead to different patterns of encoding motion events, described in the following.

2.1 Motion event typology

In his 1991 paper, Talmy develops a typology of motion event encodings, sorting the languages of the world into two types depending on the morphosyntactic element in which they express the Path component: verb-framed and satellite-framed languages (Talmy 1991: 486–487). Verb-framed languages, such as Romance languages, express the Path component in the main verb of the sentence.

(2) Spanish (Talmy 1985: 69)

La botella salió de la cueva flotando.
The bottle went.out of the cave floating
Path Manner

Other languages, such as English or German, express the Path component in a so-called satellite, a term defined by Talmy (1985: 102) as “immediate constituents of a verb root other than inflections, auxiliaries, or nominal arguments, [related] to the verb root as periphery (or modifiers) to a head”. These languages are therefore named satellite-framed languages. The English counterpart to (2) can be seen in (3).

(3) The bottle floated out of the cave.
Manner Path

Slobin (2004) revises this binary typology by adding a third type, equipollently-framed languages, suited to describe languages with serial verb constructions where both Path and Manner can be expressed in a verb, as illustrated by the Mandarin Chinese example in (4).

(4) Mandarin Chinese

Hǎiōu cóng dòng lí fēi chū.
seagull from hole in fly exit
Manner Path

‘The seagull flew out of the hole.’

Much work has followed the papers of Talmy (1991) and Slobin (2004), classifying many different languages into the different patterns. Many of these works have used the so-called “Frog Stories” also used in Slobin (2004). In sections 2.3 and 2.4, a short overview will be given of the work on motion event encoding in Kwa languages as well as French, which have been relevant to the formation of Haitian Creole.

2.2 Manner salience

Slobin (2004) takes a more detailed look at the expression of Manner, that is to say the frequency with which it is expressed in different languages. To this end, he compares the encoding of one certain event in the Frog Stories, namely an owl flying out of a knothole (Slobin 2004: 224–225). In the verb-framed languages Spanish, French, Italian, Turkish and Hebrew, virtually no Manner verbs are used, see (5) for French.

- (5) French (Slobin 2004: 224)
 D'un trou de l'arbre sort un hibou.
 from.a hole in the.tree exits an owl
 Path

Between different satellite-framed languages, more variation can be found regarding the expression of Manner. In German and English, Manner verbs are not used very frequently (only by about 17–32% of the speakers) to express the motion event in question. This is due to the fact that often deictic verbs are used with Path satellites, as in the German example in (6).

- (6) German
 Aus dem Astloch kommt eine Eule raus.
 from ART.DEF.DAT knothole comes ART.INDEF owl out
 Path

In the equipollently-framed languages Mandarin Chinese and Thai, Manner is expressed more frequently (by 40% of the Mandarin and 59% of the Thai speakers). In the SF language Russian, the Manner component of this event is expressed by 100% of the speakers. In all these cases, either a deictic (*pri-letet* 'fly here') or a Path-prefix (*vy-letet* 'fly out') is added to the Manner verb *letet* 'to fly'.

Slobin comes to the conclusion that the frequency with which the Manner component is expressed depends on the language type as well as the morphosyntactic possibilities to encode Manner. He proposes to align languages along a scale of Manner salience, where languages expressing Manner in the main verb typically have a high Manner salience, whereas languages where Manner is subordinate to Path typically have low Manner salience (Slobin 2004: 250).

2.3 Motion events in French

As already mentioned above, French is classified as a verb-framed language, encoding Path in the main verb and Manner in a gerund.

- (7) French
 Elle entrait à la maison en.courant.
 3SG.F entered PREP ART.DEF house run.GER
 Path Manner

An extensive study of motion event encoding in French can be found in Pourcel & Kopecka (2005). They analyze a total of 1800 written and 594 oral descriptions of motion events and, on this basis, describe five different patterns frequently found in French (Pourcel & Kopecka 2005: 145–149). The most frequent

is the verb-framed type, as already shown in (7). Another frequent pattern is the coordination of two verb phrases, one containing a Manner verb and the other containing a Path verb:¹

- (8) French (Pourcel & Kopecka 2005: 145)
Il court dans une rue puis rentre dans une maison.
He runs on a street then enters into a house
Manner Path

The authors find a third pattern which they call “reverse verb-framed pattern” because it is structurally identical with a verb-framed pattern but Manner and Path “switch places”, so that the Manner component is expressed in the main verb and the Path component in a gerund, see (9).

- (9) French (Pourcel & Kopecka 2005: 145)
Il court en traversant la rue.
He runs crossing the street
Manner Path

The fourth type, in which Manner is expressed in the verb and Path in a PP, is also called reverse verb-framed pattern by the authors. This fourth type can also be described as a satellite-framed construction, see (10).

- (10) French (Pourcel & Kopecka 2005: 145)
Il court dans le jardin.
He runs into the garden
Manner Path

The fifth pattern is a hybrid type because the verbs here encode Path as well as Manner. There are two subtypes to this pattern: In the first, both elements are expressed in the verb, as in (11); in the second, Path is expressed in an incorporated prefix of the verb, as in (12).

- (11) French (Pourcel & Kopecka 2005: 146)
Marc a plongé dans le lac.
Marc dived into the lake
Manner.Path

¹Motion events expressed in a single phrase are the main interest of the study, but because the coordinated pattern is so frequent in the data, it is nevertheless listed here.

- (12) French (Pourcel & Kopecka 2005: 149)
 L’oiseau s’est en-volé du nid.
 The bird has away-flown from the nest
 Path-Manner

Besides the description of these five patterns, the authors show by using acceptability judgments that in French it is dispreferred to express the Manner component of a motion event as long as it is inferable from the context. Only when the Manner of motion is not typical for the Figure or Ground of the event, it is acceptable to express it (Pourcel & Kopecka 2005: 148). This finding is in line with the observation of Berthele (2013) that Manner is seldomly expressed in French motion event encodings.

2.4 Motion events in Kwa languages

Kwa languages form part of the Niger-Congo language family, members of which show a general tendency to lexicalize Path in verbs, such as *enter*, *pass*, or *ascend* (Schaefer & Gaines 1997: 200–202). As for the expression of Manner, much variation is found between the members of this language family (Schaefer & Gaines 1997: 209).

A more detailed study on the expression of motion events has been carried out for two different Kwa languages, viz. Ewe (Ameka & Essegbey 2013) and Fon (Lambert-Brétière 2009).

In Ewe, serial verb constructions combining a Path verb and a Manner verb can be used to express motion events, see (13).

- (13) Ewe (Ameka & Essegbey 2013: 24)
 Devi-a tá yi xɔ-a me.
 child-DEF crawl go room-DEF in
 ‘The child crawled into the room.’

It is possible to combine a Manner verb with more than one Path verb, each indicating movement in respect to a different ground object, see (14).

- (14) Ewe (Ameka & Essegbey 2013: 30–31)
 Kofi tá tó ve-a me do yi kpó-á dzi.
 Kofi crawl pass ditch-DEF in exit go hill-DEF top
 ‘Kofi crawled through the ditch and emerged at the top of the hill.’

In Fon, motion events can also be expressed using verb serialization, as in (15).

- (15) Fon (Lambert-Brétière 2009: 14)
xèví ò zòn gbò tá nǔ é
bird DEF fly pass head for 3SG
‘The bird flew over his head.’

As in Ewe, a Manner verb can be combined with more than one Path verb, as in (16).

- (16) Fon (Lambert-Brétière 2009: 22)
Cùkú ó lǒn tón sín xò mè gbòn flété ó nù.
dog DEF jump exit from room in pass window DEF edge
‘The dog jumped out of the room through the window.’

Available for motion event verb serialization is a closed class of ten Path verbs (Lambert-Brétière 2009: 9). All of these can also be used outside of verb serializations, but not all Path verbs are available for serialization, like e.g. *xá* ‘go.up’ (Lambert-Brétière 2009: 16). Similarly, not all Manner verbs are available for serialization, see (17).

- (17) Fon (Lambert-Brétière 2009: 15)
* yě dǔ-wè tón sín xwé ó mè
3PL move-dance exit from house DEF in
‘They danced out of the house.’

Whereas Ameka & Essegbey (2013: 36) classify Manner-Path verb serializations in Ewe as equipollently-framed constructions, Lambert-Brétière (2009: 19) argues that the Fon Manner-Path verb serializations are satellite-framed constructions with the Path verbs acting as satellites. She reaches that conclusion because certain inflectional elements can only appear in front of the Manner verb, which marks them, in her point of view, as the main verb of the sentence.

In fact, the question how Manner-Path verb serializations should be classified in the typology described above is controversial. It depends mainly on the question whether the verbs are co- or subordinated. The discussion of this problem is outside of the scope of the present study. More details on the topic can be found in Talmy (2009).

2.5 Motion events in Haitian Creole

To my knowledge, no study has aimed at investigating the expression of motion events in Haitian Creole² until now. Nonetheless, some insights can be obtained from the literature on Haitian Creole. The language possesses an inventory of Path verbs, many of French origin, like in the example in (18).

- (18) (Fattier 2013: 203)
Dlo antre anndan kay.
water enter LOC house
'Water came into the house.'

Besides that, of all French-based Creole languages, Haitian Creole is the one that exhibits the most serial verb constructions (Mutz 2017: 44). Many of those constructions found in the literature do not express motion events, but a few examples of Manner-Path verb serializations can be found, such as the one in (19).

- (19) (Valdman 2015: 244)
Tidjo kouri ale lakay li.
Tidjo run go home POSS.PRON
'Tidjo ran over to his house.'

The dissertation on Haitian Creole verb serialization by Bucheli Berger (2009) does not offer examples of Manner-Path verb serialization, but lists the possible combinations of Manner and Path verbs, a shortened version of which is reproduced here in Table 1 (on the following page).³

3 Study design

The present study investigates the expression of motion events as presented above in Haitian Creole. The main purpose is to describe the morphosyntactic elements used to express the components Manner and Path and the preferences

²In the following, all examples are from Haitian Creole, so this will not be indicated in the rest of the paper.

³Her results are derived from online research. Marked as possible are those combinations for which she could find examples online. If a combination is not marked as possible, this does not necessarily mean that it is impossible but simply that the author could not find an example for it during her research. No acceptability study was carried out. An anonymous reviewer of this paper notes that some combinations, especially the ones with *tonbe*, sound strange to them.

Table 1: Manner of Motion V1 + Path of Motion V2 in Haitian Creole after Bucheli Berger (2009: 202)

	<i>al(e)</i> 'go'	<i>vin(i)</i> 'come'	<i>sôt(i)/</i> <i>sot(i)</i> 'go out'	<i>antre</i> 'go in'	<i>rive</i> 'arrive'	<i>monte/</i> <i>moute</i> 'go up'	<i>desann</i> 'go down'	<i>(re-) tounen</i> 'come back'
<i>kouri</i> 'run'	+	+	+	+	+	+	+	+
<i>mache</i> 'march'	+	+	+	+	+			
<i>naje</i> 'swim'			+					
<i>woule</i> 'roll'	+	+	+				+	
<i>koule</i> 'flow'			+				+	
<i>vole</i> 'fly'	+	+	+				+	
<i>glise</i> 'glide'			+				+	
<i>tonbè</i> 'fall'			+		+			

of their use. For this purpose, four Haitian Creole speakers living in Berlin, Germany, took part in interviews that consisted of three parts: an interview on their habits of language use and attitudes towards all their languages, a narration of a picture story, and descriptions of single pictures representing different motion events that were drawn by the author of this study. The entire interviews were held in Creole. One of the participants, P1, helped realize the other three interviews as well as transcribe and translate the language data recorded. She will henceforth be referred to as the main participant. More information on participants and tasks is given in the following sections.

3.1 Participants

The four participants were aged between 34 and 56. P1 is female; P2, P3 and P4 are male. P1 is a B.A. student, P2 is a mechanical engineer, P3 is a salesperson and photographer and P4 is a political scientist and educator in development cooperation. They were all born in Haiti and completed most of their education there. P1, P2 and P3 come from the area of Port-au-Prince, P4 moved there from the North of the country when he was ten years old. All four emigrated between 20 and 30 years of age. P2 and P4 regularly work in Haiti. The four participants all speak Haitian Creole, French, German, Spanish, and English. They learned Haitian Creole as their first language from their parents and later learned French in school. They received education almost entirely in French; only P1 had Creole language classes for one year. All four report they are able to converse fluently in Creole but have problems with writing, as they have never learned a norm. As the

four participants all live in Germany, they speak German on a daily basis.⁴ P1, P3 and P4 report that they speak French often, mostly with their family, especially with their children. P4 also speaks French (as well as Creole) at work. Creole is spoken with friends and family in Haiti and abroad, e.g. with their parents and siblings. P2 is the only participant that reports that he speaks Creole often, mainly with his children but also the rest of the family, as well as when working in Haiti. He is also the only one to name Creole as the language he finds most elegant; for the three others that language is French. When asked what the Creole language means to them, all four replied that it is an important part of their identity and their origin. P3 and P4 also say that they feel that Creole is the most important one of their languages.

3.2 Tasks

There were two tasks aiming at eliciting motion events, the narration of a picture story and the description of single pictures drawn for the purpose of this study. The picture story selected was *Die Geschichte vom Vogel* ('The bird story') (from Rettich & Rettich 1972).⁵ Even though the Frog Stories have been used to elicit motion events in many previous studies, they were not used here, first because they were considered difficult to narrate by the author of this study and her supervisor, and second because many of the Frog Story pictures do not contain motion events. The bird story is about a bird that flies out of its cage and then out of its house. Outside, he meets different animals that all chase him away. Finally, he flies back to his house and into his cage. The story was chosen because it contains many different motion events which could help determine how frequent the Manner component would be expressed in order to investigate the Manner salience of Haitian Creole.

The description of single pictures aimed at exploring the morphosyntactic elements that could be used to express different Manner-Path combinations. Therefore, seven Manner elements (*run, swim, fly, jump, crawl, dance, roll*) and ten Path elements (*out, away, to, into, up, down, along, past, after somebody, through*) were used to create a total of 48 motion events, as in (20), which were then portrayed in simple pictures by the author of this study.

⁴Of course, the fact that the four participants all live in a non-Creole-speaking country and use other languages on a daily basis could influence the Creole they speak causing it to be different from the Creole spoken in Haiti. Because the present study was carried out as an MA thesis, getting fieldwork data was not possible. Possible contact phenomena will not be investigated in the present study, but this has to be kept in mind when interpreting the results.

⁵Unfortunately, the image of the picture story cannot be reproduced here due to copyright reasons.

- (20) He runs out of the burning house.
Manner Path



Figure 1: Depiction of the motion event ‘to run out of’

The combination of the 17 motion event components would have yielded more than 48 events, but it was decided not to overwhelm the participants with too many pictures. The 48 drawings were divided into two groups of 24, which were presented to two participants each. When dividing the pictures, the different Manner and Path components were divided as equally as possible between the two groups. Within the two groups, the pictures were arranged in such a way that two following pictures never contained a component already depicted in the previous picture. During the interview, the participants were told to describe what the person in the picture was doing.

3.3 Data analysis

The interviews were transcribed and translated into German by the main participant. Transcriptions and translations were later checked by the author of this paper and revised together.

At the beginning of the data analysis, the number of sentences was counted for the picture story narrations. Every unit containing a subject and (at least)

one verb was counted as one sentence. Coordinated clauses were counted as two sentences, but subordinated clauses like complement clauses, relative clauses, causal clauses, temporal clauses and the like were counted as part of the matrix clause.

P1's narration contains 26 sentences, P2's contains 46 sentences, P3's contains 55, and P4's narration contains 42 sentences. In total, 169 sentences were elicited.

After counting the number of sentences, the number of motion events encoded was determined. Every sentence expressing directional motion was analyzed as a motion event encoding.

P1's narration contains 18, P2's 20, P3's 29, and P4's 27 motion event encodings. Altogether, the four participants encoded 94 motion events in their picture story narrations. With a total of 169 sentences, more than 50% of the sentences contained a motion event encoding.

The motion events were then sorted by means of the morphosyntactic structure they used to encode different motion event components. They were sorted into six different categories: Path verbs only, see (21), Path verbs with Ground-PP/NP, see (22), Manner verbs only, see (23), Manner verbs with path elements, see (24), serial verb constructions, see (25), and motion events without a motion verb, see (26). The remaining cases were classified as "Other".

- (21) Li rantre.
3SG enter.again
'He goes back in.'
- (22) Epi l antre nan kay la.
and 3SG enter LOC house DEF
'He enters the house.'
- (23) E zwazo a kouri.
and bird DEF RUN
'And the bird runs/flies fast.'
- (24) [...] zwazo a vole sou do yon erison.
bird DEF fly LOC back INDEF hedgehog
'The bird flies onto the back of a hedgehog.'
- (25) Li kouri retounen nan kay [kote li te ye a.]
3SG run return LOC house [REL.PRON 3SG PST COP DEF]
'He goes back into the house where he was before.'

- (26) Epi li kraze rak.
then 3SG destroy forest
'Then he beats loose.'

The results of the analysis will be given in the following section.

The first step of the picture description analysis was to determine the number of descriptions. As 48 pictures were described by two participants each, 96 descriptions should have been elicited, but as one of the participants failed to describe two of the pictures, only 94 descriptions were elicited. Some of the descriptions consist of a simple sentence, whereas others consist of a complex sentence or even more than one sentence. A total of 119 sentences were elicited in both tasks.

If more than one sentence was used for the description of a picture, they were counted and analyzed separately. The same holds for complex sentences if they contained more than one motion event, e.g. a sequence of two relative clauses, as in (27), or sentences with *pou* 'in order to', as in (28).

- (27) Yon zwazo k ap vole k ap pase bò kot yon
INDEF bird REL.PRON PROG fly REL.PRON PROG PASS beside side INDEF
pyebwa.
tree

'A bird which is flying, who is passing next to a tree.'⁶

- (28) Yon mesye k ap naje sòti nan plaj pou l ale bò
INDEF man REL.PRON PROG swim exit LOC beach for 3SG go beside
rivaj.
coast

'A man who is swimming away from the beach in order to swim to the coast.'

Sentences which did not express motion (13 of 119) were not analyzed.

In a few cases, modal verbs were used, see (29) and (30). These were ignored for the analysis and the event encodings of the motions were treated as if they did not contain a modal verb.

⁶Mostly P2, but also P4, described several of the pictures with utterances of the form NP + relative clause. Even though these do not constitute regular sentences, it is possible to analyze them as elliptic versions of sentences like *This is [NP] who is moving* which are also found in some descriptions. They were therefore included in the analysis.

- (29) Yon gason ki vle monte sou yon tab.
 INDEF boy REL.PRON want ascend LOC INDEF table
 ‘A boy who wants to go up onto a table.’
- (30) Yon mesye ki dwe travèse dyagonal yon chanm.
 INDEF man REL.PRON must cross diagonal INDEF room
 ‘A man who has to cross a room diagonally.’

The motion events expressed in the picture descriptions were then sorted into eight categories, seven of which are equivalent to those for the picture story. A new category was established for this part of the data: Manner verbs with Ground elements, as exemplified in (31).

- (31) Yon moun k ap rale kote yon mi.
 INDEF person REL.PRON PROG crawl beside INDEF wall
 ‘A person who is crawling next to a wall.’

The results of the analysis are given in the following section.

4 Results

An overview of the results is given in Table 2. In the following subsections, the results are discussed in detail.

Table 2: Motion event expressions in picture story narrations and single picture descriptions

	Picture story		Single pictures		Total	
Path verb only	10	10.6%	4	3.4%	14	6.6%
Path verb + ground PP/NP	25	26.6%	29	24.4%	54	25.4%
Manner verb only	18	19.1%	19	16.0%	37	17.4%
Manner verb + ground	0	0.0%	13	10.9%	13	6.1%
manner verb + path element	2	2.1%	5	4.2%	7	3.2%
SVC	16	17.0%	36	30.3%	52	24.4%
Motion event without motion verb	7	7.4%	1	0.8%	8	3.8%
Other	16	17.0%	12	10.1%	28	13.1%
Total	94		119		213	

Path verbs only were used ten times in the picture story narrations (10.6% of all occurrences). In most cases, the Ground was mentioned in the preceding or following context but not in the same clause, see (32) and (33).

- (32) E li ouvri pòt kalòj la pou li kapab sòti.
and 3SG open door cage DEF for 3SG able.to exit
'And she opens the door of the cage so that he can go out.'
- (33) Epi zwazo a tounen. L al nan menm kay la [...]
and bird DEF return 3SG go LOC same house DEF
'And he returns. He goes into the same house.'

In one case, no ground is mentioned at all, see (34).

- (34) Papiyon an ale.
butterfly DEF go
'The butterfly goes/flies away.'

At this point, the picture story shows a butterfly flying away from the bird. Therefore, *ale* seems to express not simply 'go' but 'go away' here.

In the single picture descriptions, Path verbs only occurred in four motion events (3.4%). As in the narrations, the Ground was usually mentioned in the context, see (35).

- (35) Yon moun k ap rale kote yon mi. L ap pase
INDEF person beside REL.PRON PROG crawl beside INDEF wall 3SG PROG
[...]
pass
'A person is crawling next to a wall. He is passing [it]...'

Again, there was one case where no Ground was mentioned at all, again with the verb *ale*, which seems to mean 'go away' (36).

- (36) Yon zwazo ki sòti nan kalòj pou ale.
INDEF bird REL.PRON exit LOC cage for go
'A bird who leaves the cage in order to go/fly away.'

4.1 Path verb + Ground NP/PP

The most frequent pattern used to express motion events in the picture story is a Path verb with a Ground NP or PP, which was used in 25 cases (26.6%). The verbs *antre* 'enter', *pase* 'pass', *atèri* 'land', *tonbe* 'fall', *ale* 'go', *poze* 'sit down', and *sòti* 'go out' were used with PPs (37).

3 The expression of motion events in Haitian Creole

- (37) Li atèri sou flè a.
3SG land LOC flower DEF
'He lands on the flower.'

The verbs *fwenn* 'reach', *suiv* 'follow', *kite* 'leave', *tounen* 'come back' were used with object NPs (38).

- (38) Li kite do erison an.
3SG leave back hedgehog DEF
'He leaves the back of the hedgehog.'

The verb *rive* 'arrive' was used with a PP three times (by P2 and P3) and with an object NP once (by P1), see (39) and (40). Because of the small number of occurrences, nothing can be said about whether this is simply due to individual preferences.

- (39) Lè l rive sou pyebwa [...]
when 3SG arrive on tree
'When he arrives on the tree...'
- (40) Zwazo a rive lakay li.
bird DEF arrive home POSS.PRON
'The bird arrives at his house.'

In the single picture descriptions, Path verbs with Ground NPs or PPs present the second most frequent pattern with 29 occurrences (24.4%).

Used with an object NP were the verbs *depase*, *desann*, and *kite* (41).

- (41) Tidjo kite lekòl la.
Tidjo leave school DEF
'Tidjo leaves the school.'

The verbs *antre*, *rantré*, *pase* and *al(e)* were used with PPs. *Antre* and *rantré* were used with *nan* 'into', *al(e)* with *bò* 'next to' and *nan direksyon* 'in the direction of', and *pase* also with *bò*, see (42).

- (42) L ap pase bò yon kay.
3SG PROG pass beside INDEF house
'He is passing a house.'

The verbs *monte* and *sòti* were used with both NPs and PPs. *Sòti* was used with three different prepositions, *nan*, *sou* and *a travè*. See (43) for an example with *a travè*, and (44) for the use with an NP.

- (43) Sa se yon moun ki sòti a travè yon fenèt [...]
DEM COP INDEF person REL.PRON exit through INDEF window
‘That is a person who leaves through a window.’
- (44) Yon timoun ki sòti lekòl.
INDEF child REL.PRON exit school
‘A child that leaves school.’

As there are only a few occurrences for each verb, often just one but six at the most, it remains unclear whether the use with NP or PP attested here is a general preference of the verb or whether all verbs can appear with both.

4.2 Manner verb only

A Manner verb alone cannot, strictly speaking, express a motion event as it is defined above, but because they occur so frequently in both picture story narrations and single picture descriptions, they are taken into account here.

In the picture story task, in the 18 cases counted for this category (19.1%), only three different Manner verbs were used, *vole* ‘fly’, *kouri* ‘run’, and *mache* ‘walk’. The last of the three is used only once where a hedgehog continues walking after the bird has landed on his back. The most frequently used of these verbs is *vole*. This is not surprising when taking into account that the story is about a bird and also features other flying animals like owls or butterflies. *Kouri* was used six times to describe the motion of the bird, see (45).

- (45) Lè chwèt la kouri dèyè zwazo a, sa k pase, zwazo
when owl DEF run behind bird DEF DEF REL.PRON happen bird
a kouri.
DEF run
‘When the owl flies behind/after the bird, the bird runs/flies away fast.’

As the story shows several instances of an animal chasing another animal (mostly the little bird) away, the cases where *vole* and *kouri* are used alone always describe a situation where the animal flees. Apparently, in these cases, a directed motion away from the place of action seems to be described. The Path

3 The expression of motion events in Haitian Creole

‘away’ seems to be inferable from the context and is therefore left out. *Kouri* obviously does not express the Manner ‘run’ in these cases, but rather an accelerated manner of movement.

In the single picture descriptions, Manner verbs were used alone in 19 cases (16% of all occurrences). In seven of these, no further information on the motion event was given, see (46).

- (46) Yon mesye ki ap naje.
 INDEF man REL.PRON PROG swim
 ‘A man that is swimming.’

These cases do not express a motion event as it is understood here, but an activity.

In the remaining twelve cases, more information on the motion event is given in the preceding or the following context. In five cases, the manner verb is followed by a construction with *pou* ‘to’ in which Path is expressed, see (47).

- (47) Tidjo ap naje pou l depase lòt la.
 Tidjo PROG swim for 3SG pass other DEF
 ‘Tidjo is swimming in order to pass the other.’

In three cases, information on the Path is given in the preceding or following sentence, see (48) and (49). As the translations show, it is possible to interpret these cases as single complex motion events.

- (48) Tidjo ap naje. Li kite il la.
 Tidjo PROG swim 3SG leave island DEF
 ‘Tidjo is swimming. He leaves the island./Tidjo is swimming away from the island.’

- (49) Tidjo antre nan kay. [...] L ap danse.
 Tidjo enter LOC house 3SG PROG dance
 ‘Tidjo enters the house. He is dancing./Tidjo is dancing into the house.’

In two cases, Path and Manner are expressed in two precedent relative clauses, see (50). Again, it is possible to interpret this as a single complex motion event.

- (50) Yon zwazo k ap vole k ap pase bò yon mi.
 INDEF bird REL.PRON PROG fly REL.PRON PROG pass beside INDEF wall
 ‘A bird who is flying, who is passing beside a wall./A bird who is flying past a wall.’

In one case, two main clauses, one containing a Manner and the other a Path verb, are combined with the conjunction *pandan* ‘while’, see (51).

- (51) L ap danse pandan l ap monte mach eskalye a.
3SG PROG dance while 3SG PROG ascend step stairs DEF
‘He dances while he goes up the stairs./He dances up the stairs.’

All these examples seem to represent complex motion events whose components were not expressed in a single clause, meaning that they were not conflated into one event.

4.3 Manner verb + Ground element

Similar to the previous category, Manner verbs combined with a Ground element in the same clause do not constitute motion events as defined by Talmy (1985), because the Path element obligatory for motion events is not encoded. Such a pattern does not occur in the picture story narrations, but there are 13 such occurrences in the single picture descriptions (10.9%). Six of the seven manner verbs tested in this task were used in this pattern, see (52) for an example with *naje* ‘swim’.

- (52) Yon moun k ap naje nan lanmè.
INDEF person REL.PRON PROG swim LOC sea
‘A person who is swimming in the sea.’

4.4 Manner verb + Path element

The least frequent of the patterns is the use of a Manner verb in combination with a Path element. Such cases constitute instances of the satellite pattern described above.

In the picture story descriptions, two such cases (2.1%) occurred with the verb *vole*, once with the preposition *sou* ‘onto’, see (53), and once with *deyò* ‘out of’, see (54).

- (53) [...] zwazo a vole sou do yon erison.
bird DEF fly LOC back INDEF hedgehog
‘A bird flies onto the back of a hedgehog.’
- (54) Li vole deyò.
3SG fly outside
‘He flies outside.’

In the single picture descriptions, this satellite-framed pattern is used in five cases (4.2%), three of them with the verb *vole*, see (55), one with *naje*, see (56), and one with *woule* ‘roll’.

- (55) Yon zwazo k ap vole a travè nyaj yo.
 INDEF bird REL.PRON PROG fly through cloud PL
 ‘A bird that is flying through the clouds.’
- (56) Tidjo ap naje [...] sou lòt bò lak la.
 Tidjo PROG swim LOC other side lake DEF
 ‘Tidjo is swimming to the other side of the lake.’

4.5 Serial verb constructions

Serial verb constructions are used frequently in both the picture story narrations and the single picture descriptions. As for the story narrations, they present the second most frequent pattern with 16 occurrences (17%). Most of these consist of a serialization of two Path verbs: *sòti kite* ‘leave go away’, *al(e) poze* ‘go sit down’, *vin poze* ‘come sit down’, *al tonbe* ‘go fall’ and *al antre* ‘go enter’. In two cases, a Manner verb is followed by a Path verb: *vole poze* ‘fly sit down’ and *kouri retounen* ‘run return’. One single verb serialization consists of three verbs: *leve pran kouri* ‘get up take run’. *Pran* most probably acts as an aspectual marker here, encoding inchoativity (see also Valdman 2015: 231). Besides the fact that this serial verb construction consists of three verbs, it is also different from the others because the order of the Path and Manner verb is inverted here in regard to the other cases: the Path verb is the first, the Manner the last verb of the serialization.

The different serial verb constructions occur either with a Ground NP, a Ground PP, or with no Ground element in the same clause. In the last case, Ground is mentioned in the surrounding context.

In four serial verb constructions, *vole poze*, *vin poze*, *al(e) poze* and *al tonbe*, the same kind of action is expressed: the bird flying to a certain place and coming to rest. It is not clear whether these are sequential or simultaneous verb serializations, the first meaning ‘he flies and then sits down’ and the latter meaning ‘he flies onto [the tree]’. The same problem exists with *al antre* ‘go enter’, where it is unclear whether it is said that the bird first goes and then enters or whether he ‘goes into’.

The case of *sòti kite* is described by Bucheli Berger (2009: 79–81) as a serialization of two verbs that are close to synonyms and which she interprets as a simultaneous serial verb construction expressing one simple motion. It is also

possible that the verbs have different meanings here, expressing the Paths ‘out of’ and ‘away’, which would make this a sequential serial verb construction.

In *kouri retounen*, which once again describes the motion of the verb, *kouri* cannot be interpreted as expressing the Manner ‘to run’, but rather a fast way of moving.

Leve pran kouri is probably a sequential serial verb construction, expressing that the bird first gets up and then flies away fast, where the Path ‘away’ is left unexpressed and to be inferred from the context.

In the single picture descriptions, serial verb constructions were the pattern most frequently used by the participants to encode a motion event. With 36 occurrences in total (30.3% of all occurrences), 31 different verb combinations can be described. The different internal structures of these serial verb constructions are summarized in Table 3.

Table 3: Frequency of different serial verb constructions in the single picture descriptions

Type of SVC	manner-path	manner-manner	path-path	path-Other	3 verbs		
					MMP	PPP	PMP
Frequency	19	2	1	6	1	1	1

By far the most common verb serializations are those consisting of a Manner and a Path verb. Six different Manner verbs were used in such constructions: *kouri* ‘run’, *naje* ‘swim’, *vole* ‘fly’, *rale* ‘crawl’, *woule* ‘roll’, and *glise* ‘glide/slide’ (used once instead of ‘roll’). Two of the Manner verbs investigated were not used in Manner-Path serializations, *sote* ‘jump’ and *danse* ‘dance’. Table 4 shows the different Path verbs that these Manner verbs were used with. Nothing can be said about the possibility to form serial verb constructions other than the ones attested in the data of this study.⁷

These Manner-Path serial verbs also occur both with NPs and PPs, see (57) and (58).

- (57) Tipyè rale monte mòn nan.
 Tipyè crawl ascend mountain DEF
 ‘Tipyè crawls up the mountain.’

⁷An anonymous reviewer of this paper notes that more combinations than the ones attested in this study should be possible, especially with *rale*. Also, all manner verbs should be able to combine with *ale*.

3 The expression of motion events in Haitian Creole

Table 4: MANNER-PATH-SVC in the single picture descriptions

	<i>sòti</i>	<i>kite</i>	<i>antre</i>	<i>monte</i>	<i>desann</i>	<i>pase</i>	<i>travèse</i>	<i>ale</i>	<i>poze</i>	<i>suiv</i>
<i>kouri</i>	+	+	+	+	+	+	+			
<i>naje</i>	+									
<i>vole</i>	+		+					+	+	+
<i>rale</i>	+			+						
<i>woule</i>					+	+				
<i>glise</i>	+				+					

- (58) *Zwazo a vole antre nan kizin nan.*
 bird DEF fly enter LOC kitchen DEF
 ‘The bird flies into the kitchen.’

In one case, no ground element is given in the same clause, see (59). As the Path verb is *ale* in this case, the Path can probably again be interpreted as ‘away’ in this case.

- (59) *Yon zwazo [...] k ap vole ale.*
 INDEF bird REL.PRON PROG fly go
 ‘A bird that is flying away.’

In two cases, two Manner verbs were serialized, see (60) and (61). In (60), it is obvious that again *kouri* cannot be interpreted as ‘run’, but most probably means that the motion takes place fast. In (61), it is unclear what the exact meaning of *vole* is. It could possibly be interpreted as having a figurative meaning expressing something like jumping into the water in a high arc. This hypothesis cannot be tested here.⁸

- (60) *Tipyè kouri naje dèyè yon lòt.*
 Tipyè run swim behind INDEF other
 ‘Tipye swims fast behind/after another.’
- (61) *Li vole sote nan dlo a.*
 3SG fly jump LOC water DEF
 ‘He jumps into the water (in a high arc).’

⁸An anonymous reviewer suggests that the two verbs act like synonyms here.

The only Path-Path serialization in the single picture descriptions is *sòti kite*, which also occurred in the picture story narrations.

In six cases, the participants used serial verb constructions consisting of a Path and a non-motion verb, like in (62).

- (62) Yon moun ki ap antre kache nan yon gwòt.
INDEF person REL.PRON PROG enter hide INDEF LOC cave
'A person who is going into a cave in order to hide.'

In three of these six cases, the verb *ale* was used, but cannot be interpreted as a Path verb, see (63). It could be interpreted as an analytical future, but such a structure with this function has not been described for Haitian Creole (see for example, Valdman 2015, DeGraff 2007). This structure cannot be further analyzed at this point.

- (63) [Sa se Johana k ap kouri] pou l al pran bis la nan stasyon an.
for 3SG go take bus DEF LOC station DEF
'[That's Johanna who is running] in order to take a bus at the station.'

In three cases, three motion verbs are serialized. The first of them is a Manner-Manner-Path serialization, see (64). As this sentence is about a swimming person, *kouri* probably once again express a fast motion.

- (64) Li kouri naje kite il la.
3SG run swim leave island DEF
'He swims away from the island fast.'

The second of these serializations consists of three Path verbs (65). *Rive* and *jwenn* are close to synonyms and are therefore interpreted as expressing the same meaning here. Together with *avanse* 'move forward' they probably present a sequential serial verb construction.

- (65) Li avanse rive jwenn demwazèl la.
3SG advance arrive reach lady DEF
'He advances towards and reaches the lady.'

In the last of the three cases, we find a Path-Manner-Path verb serialization, see (66). Interestingly, V1 and V3 are the same verb, *sòti*. As this is the only case where we find such a structure in the data, nothing can be said about whether this is a common pattern of serial verb constructions in Haitian Creole.

- (66) Boul la sòti woule sòti nan bwat katon.
 ball DEF exit roll exit LOC box cardboard
 ‘The ball rolls out of the cardboard box.’

Even though many of the serial verb constructions elicited in this study are single cases that need to be described separately, some patterns can also be found. One frequent strategy to express complex motion events in Haitian Creole is the use of a Manner-Path verb serialization, which in a few cases also occurred in serializations of three verbs. In other cases, two Path verbs or a Path and a non-motion verb were serialized to encode a motion event depicted in one of the drawings.

4.6 Motion events without a motion verb

In the picture story narrations, seven cases occurred where a motion event was expressed without a motion verb (7.4%). Five of these were uttered by P3, when he used the idiom *kraze rak*, which has the meaning ‘to beat loose’. The two other cases were uttered by P2 using *jwenn direksyon* and *pran direksyon* to say that the bird was going into a certain direction, see (67).

- (67) [...] kounye a la li jwenn direksyon fenèt la
 moment DEF DEM 3SG reach direction window DEF
 ‘Now he takes the direction of the window.’

In the single picture descriptions, there is only one case where a motion event is expressed without a motion verb (0.8%), also using *direksyon*, see (68).

- (68) Li pran nan direksyon machin nan.
 3SG take LOC direction car DEF
 ‘He takes the direction of the car.’

4.7 Other

The remaining encodings of motion events had to be sorted into a separate category because it was not possible to analyze them as any of the other categories.

For the picture story narrations, all of the 16 cases sorted into this category (17%) use the preposition *dèyè* ‘behind’ or ‘after’, 13 with *kouri*, see (69), and 3 with *pati*, see (70). In all of these cases, one animal is chasing another.

- (69) Chwèt la kouri dèyè zwazo a pou l pran l.
owl DEF run behind bird DEF 3SG take 3SG
'The owl flies behind/after the bird in order to catch him.'
- (70) Koukou a pati dèyè zwazo a.
cuckoo DEF leave behind bird DEF
'The cuckoo flies behind/after the bird.'

The problem here is the preposition *dèyè*: If it expresses 'behind', the PP can be analyzed to encode the Ground and locate the place where the motion is taking place. If it expresses 'after', it can be analyzed as expressing the Path of motion.

The main participant translated all of these cases with *chase away*. This could be the implication of flying fast behind someone. In one case, however, P4 uses *kouri dèyè* where the subject isn't moving at all. The picture shows a group of birds sitting in a tree chasing away the little bird by screaming at him (71).

- (71) Zwazo sa yo genlè pa renmen li. Yo kouri dèyè li. E lè
bird DEM PL seem NEG like 3SG PL run behind 3SG and when
sa li vole.
DEM 3SG fly
'These birds seem to not like him. They chase him away. And then he flies away.'

Another problem is the semantics of *pati*. P1 uses *kouri dèyè* and *pati dèyè* in a similar way. When the owl is chasing the little bird by flying after him, she uses *pati dèyè*. The semantics of *dèyè* are obscure here, because the Path 'away from X' is not relevant here. It is neither shown in the pictures nor expressed in the narration.

For the single picture descriptions, twelve descriptions had to be sorted into the category "Other" (10.1%). Three cases were equivalent to those in the picture story narrations where *dèyè* was used. In two other cases, hybrid verbs were used which could not be clearly identified as Manner or Path verbs as they contain both components: *plonje* 'dive into' and *tonbe* 'fall'.

In (72), a Path verb is combined with a further description of the Path.

- (72) Yon mesye ki dwe travèse dyagonal [...] yon chanm.
INDEF man REL.PRON must cross diagonal INDEF room
'A man who must cross a room diagonally.'

In (73), the Manner component is expressed in a PP, which is the only occurrence of this type.

- (73) Tidjo ap monte mòn ak kat pat.
 Tidjo PROG ascend mountain with four paws
 ‘Tidjo is crawling up the mountain.’

The four remaining occurrences contain a gerund construction, all used by the same participant, P1. In two of these cases, a structure equivalent to the French structure *V en V.GER* is used, see (74). This structure has not been described for Haitian Creole.

- (74) Li desann eskalye a an dansan.
 3SG descend stairs DEF PREP dance.GER
 ‘He/She goes down the stairs dancing.’

In the other two cases, the Path verb is followed by a gerund form of ‘to be’, *etan*, and then a full sentence consisting of subject, aspect marker and manner verb, see (75). This structure has also not been described for Haitian Creole.⁹

- (75) Tipyè antre etan l ap danse nan chanmnam.
 Tipyè enter be.GER 3SG PROG dance LOC room DEF
 ‘Tipyè is dancing into the room.’

5 Discussion

In the previous section, the morphosyntactic patterns used to express motion events in the present study were described. Most commonly used were three different structures: a Path verb with a Ground NP or DP, a serial verb construction, or a Manner verb only. Path verbs with Ground elements are verb-framed structures in the sense of Talmy (1991). In all of those cases, the Manner component of the event was not expressed. Serial verbs often consisted of a Manner and a Path verb, which could be labelled an equipollently-framed construction in the sense of Slobin (2004). Other verb serializations were also found, mostly of two Path verbs, but also combinations of Path and non-motion verbs. These are also verb-framed constructions. The third most frequent pattern is the use of a Manner verb only, with no Path element encoded in the same clause. According to Talmy’s (1985) definition, the latter is not a motion event. These cases are nevertheless taken into account here, primarily because of their relatively high

⁹Both a native speaker present at the talk at the SPCL Meeting in Tampere as well as an anonymous reviewer of this paper noted that this structure is very uncommon in Haitian Creole and most probably due to other language influence.

frequency. Besides that, it is possible that at least some of them do, contrary to Talmy's definition, express complex motion events, because the Path component is possibly left to be inferred in these cases but implicitly present. This was the case in some examples from the picture story narrations, where the participants uttered sentences like *Zwazo a kouri/vole* 'The bird flies (fast)', meaning that the bird is flying away. Most of the uses of a sole Manner verb in the single picture descriptions are probably due to the fact that the depicted motion event was too difficult to recognize as such, as in *swimming along the coast*. In these cases, the participants simply expressed a motion activity instead of the event.

Besides these three main patterns, three further but marginal patterns were found in the data: a Manner verb with a Path element, a Path verb alone and a motion event without any motion verb. Manner verbs with Path elements constitute satellite-framed constructions in the sense of Talmy (1991), which were rare but are still attested in the present data. They occurred with few, but different Manner verbs and also different Path elements. When Path verbs were used alone, information on the Ground was usually given in the context. The expression of motion events with idioms or constructions like *pran nan direksyon* 'take a certain direction' is most probably not typical for Haitian Creole but occurs in many languages.

The different patterns described above show that Haitian Creole possesses a rich inventory of morphosyntactic structures available to express motion events. It is therefore not classified here as a language of one of the three types described above, VF, SF and EF languages. All of these three patterns are found in the Haitian Creole data, VF and EF patterns being more frequent than SF patterns.

Some problematic cases were also described above which need further investigation. In the cases where *dèyè* was used, it was not clear whether it expresses 'behind' or 'after' and it could therefore not be decided whether it constitutes a Ground or a Path element. This shows that clear semantic criteria to identify the components of motion events are needed. Such criteria could also help to further investigate hybrid verbs like *plonje* 'dive' which are said to express both Manner and Path. Two other cases were also problematic as they presented completely different structures from the ones described earlier. The structures where a gerund of a Manner verb or of the verb 'to be' were used, neither of which has been described for Haitian Creole. As the participants of this study all lived outside of Haiti and were using other languages such as German or French on a daily basis, it is possible that these structures are due to language contact, most probably with French. More research needs to be done in this area.

The second aim of this study was to investigate the Manner salience of Haitian Creole, that is the frequency with which the Manner component is encoded in

3 The expression of motion events in Haitian Creole

motion event expressions in comparison with other languages. In the data described above, Manner was expressed either in a Manner-Path verb serialization or in a satellite construction.¹⁰ In the picture story, both Manner-Path verb serializations and satellite constructions were very rare, which means that Manner was often left unexpressed. As the story was about a bird, whose Manner of motion typically is to fly, it is not necessary to encode Manner in every motion event, as it can easily be inferred. In the single picture descriptions, Manner-Path verb serializations are used in 24.7% (21 of 85) and satellite-framed constructions in 5.9% (5 of 85) of the motion event expressions. With a total of 30.6%, the frequency of Manner encodings is much higher here than in the picture story narrations. Considering the fact that all of the pictures showed a specific Manner component, this number is nevertheless rather small. Both the picture story as well as the single picture descriptions therefore indicate that Haitian Creole has low Manner salience.

In comparison to the French motion verb expressions described in the first part of the paper, some similarities and some differences can be observed. Just like French, Haitian Creole possesses a rather large inventory of Path verbs, most of which probably go back to their French counterparts. This is why VF constructions are common in both languages. However, their percentage is larger in French than in Haitian Creole, as the latter possesses another structure not available in French: verb serialization, particularly the serialization of a Manner and a Path verb. This is a feature that Haitian Creole shares with various African languages, Kwa languages in particular, which are said to have played a significant role in the formation of Haitian Creole. Just as in the Kwa languages described above, the Manner verb precedes the Path verb in the Haitian Creole motion verb serializations. Another interesting observation is the fact that in the present data, no serializations with the Manner verb *danse* 'to dance' are attested, a Manner-Path serialization which is ungrammatical in Fongbe according to Lambert-Brétière (2009). The (un)grammaticality of such serializations in Haitian Creole needs to be tested in a subsequent study. The preliminary result of the ongoing research on this question is that the morphosyntactic patterns used in Haitian Creole to express motion events seem to be a mixture of the patterns found in the languages that were relevant to its formation.

In further research, more data will be elicited. As some of the drawings used for the single picture descriptions proved to be difficult to interpret, these rep-

¹⁰As Manner verbs only and Manner verbs with ground expressions are not considered motion events as defined by Talmy (1985), they are not included here. This leaves us with 76 motions event expressions in the picture story narrations and 85 motion event expressions in the single picture descriptions.

representations of motion events need to be revised. If possible, videos showing motion events would be preferred for data elicitation. In addition, acceptability judgments will be elicited to investigate which Manner and Path elements are possible in which pattern, mostly in serial verb constructions and satellite-framed constructions.

Acknowledgements

I am grateful for the ongoing support of the supervisor of my thesis, Prof. Judith Meinschaefer. Further thanks go to the native speaker who helped me transcribe and translate the Haitian Creole data. I would also like to thank two anonymous reviewers who have provided very helpful comments on this paper. All remaining errors are my own.

References

- Ameka, Felix & James Essegbey. 2013. Serialising languages: Satellite-framed, verb-framed or neither. *Ghana Journal of Linguistics* 2(1). 19–38.
- Berthele, Raphael. 2013. Disentangling manner and path: Evidence from varieties of German and Romance. In Juliana Goschler & Anatol Stefanowitsch (eds.), *Variation and change in the coding of Motion events*, 55–75. Amsterdam: John Benjamins.
- Bucheli Berger, Claudia. 2009. *Serielle Verbkonstruktionen in Frankokreolsprachen: Beschreibung, Genesetheorie und typologischer Vergleich* (Arbeiten des Seminars für Allgemeine Sprachwissenschaft der Universität Zürich 21). Zürich: University of Zürich.
- DeGraff, Michel. 2007. Kreyòl Ayisyen, or Haitian Creole. In John A Holm & Peter Patrick (eds.), *Comparative Creole syntax: Parallel outlines of 18 Creole grammars*, 101–126. Battlebridge: Battlebridge Publications.
- Fattier, Dominique. 2013. Haitian Creole. In Susanne Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of Pidgin and Creole languages, vol II, Portuguese-based, Spanish-based, and French-based languages*, 195–204. Oxford: Oxford University Press.
- Lambert-Brétière, Renée. 2009. Serializing languages as satellite-framed. *Annual Review of Cognitive Linguistics* 7. 1–29. DOI: 10.1075/arcl.7.01lam.
- Mutz, Kathrin. 2017. Aktuelle Forschungsfragen der Kreolistik. In Peter Stein (ed.), *Kreolisch und Französisch*, 2nd edn. (Anglophone Postcolonial Studies), 53–70. Tübingen: Niemeyer.

- Pourcel, Stéphanie & Anetta Kopecka. 2005. Motion expression in French: Typological diversity. *Durham & Newcastle Working Papers in Linguistics* 11. 139–153.
- Rettich, Rolf & Margret Rettich. 1972. *Hast du Worte?* Ravensburg: Otto Maier Verlag.
- Schaefer, Ronald P. & Richard Gaines. 1997. Toward a typology of directional motion for African languages. *Studies in African Linguistics* 26(2). 193–220. DOI: 10.32473/sal.v26i2.107393.
- Slobin, Dan. 2004. The many ways to search for a frog: Linguistic typology & the expression of motion events. In Sven Strömquist & Ludo Vorhoeve (eds.), *Relating events in narrative: Typological and contextual perspectives*, vol. 2, 219–257. Mahwah, NJ: Lawrence Erlbaum Associates.
- Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Timothy Shopen (ed.), *Language typology and syntactic description: Grammatical categories and the lexicon*, vol. 3, 57–149. Cambridge: Cambridge University Press.
- Talmy, Leonard. 1991. Path to realization: A typology of event conflation. In Laurel Sutton, Christopher Johnson & Ruth Shields (eds.), *Proceedings of the seventeenth annual meeting of the Berkeley Linguistics Society*, 480–519. Berkeley: Berkeley Linguistics Society.
- Talmy, Leonard. 2009. Main verb properties and equipollent framing. In Jian-sheng Guo, Elena Lieven, Nancy Budwig, Susan Ervin-Tripp, Keiko Nakamura & Şeyda Özçalışkan (eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin*, 389–402. New York: Psychology Press.
- Valdman, Albert. 2015. *Haitian Creole: Structure, variation, status, origin*. Sheffield: Equinox.

Chapter 4

Postulating Atlantic English Pidgin/Creole as a pluriareal language: A perception study

Angela Bartens^{a,b}, Kwaku Owusu Afriyie Osei-Tutu^c &
Tamirand Nnena De Lisser^d

^aUniversity of Turku ^bUniversity of Helsinki ^cUniversity of Ghana ^dUniversity of Guyana

This paper aims at verifying the observation that many speakers of closely related Atlantic English(-lexifier) Pidgin(s) and Creole(s), or EP/C(s), can understand each other. In order to test mutual intelligibility of different EP/Cs, an electronic survey was completed by 56 participants from Ghana, Guyana, and Nigeria. In addition, 20 interviews were conducted with informants representing eight different nationalities and residing in Guyana at the time. The data were also supplemented by a reading comprehension and translation task of short passages in eight Atlantic EP/Cs conducted in Ghana and Guyana. The results of the study indicate that intelligibility exists between most (but not all) language pairs. We therefore suggest that we are dealing with a case of pluriareality rather than pluricentricity, as the habitual criteria for the latter are not (yet) fulfilled. We believe that this can be accounted for by a postcolonial understanding of language variation, allowing for the postulating of a language system with fluid borders for use in wider contexts, including writing. Nevertheless, the postcolonial aspect shall be addressed in more detail in a posterior study.

1 Introduction

It is almost taken for granted that speakers of many, albeit not all, Atlantic English-lexifier Pidgins and Creoles, can understand each other and are aware of



the linguistic similarities.¹ We aim to go beyond this impressionistic observation by using, for the time being, three distinct verification methods: an online survey, interviews, and a reading comprehension and translation task for checking the plausibility of this observation.

As far as terminology is concerned, an abundance of definitions for the terms PIDGIN and CREOLE exists, highlighting the differences between what are often conceived of as distinct types of languages as far as complexity and range of use are concerned (cf. Bakker & Matras 2013: 6). However, linguistic reality is more multifaceted and, without proposing to return to the discussion regarding the “life cycle” of pidgins (Hall 1962) or the complexity of Creoles (see, for example, McWhorter 2011), it is evident that some pidgins become the dominant or native languages of their speakers², resulting in what were initially called “expanded pidgins” and later on “pidgincreoles”, a term attributed to Philip Baker (Bakker 2008: 113). Nigerian Pidgin or Naija, one of the languages and speech communities studied here, is a case in point: widely spoken as a second language, it has a growing number of native speakers, especially in some areas such as the Warri-Sapele area in Delta State (cf. Deuber 2005: 1, 3). For our purposes, attributing language varieties to such categories is irrelevant. Rather, we used as inclusive a denomination as possible so that informants³ could identify with a large range of such languages and therefore chose “English(-lexifier) Pidgin(s) and Creole(s)”, henceforth EP/C(s).⁴

We aim at testing mutual intelligibility. Intelligibility can be defined as “the degree to which a spoken message can be understood” (Lingualinks cited by Bouwer 2007: 257), and is determined by such criteria as “rate of speech, intonation, location of pauses, topic familiarity, discourse structure, grammatical complexity, and hearer’s attitude toward the text or speaker” (ibid.). As we shall see, some of these are highly relevant for our informants. It can be argued that intelligibility is essentially about social relationships, rather than linguistic relationships between varieties (Romaine 1994: 14). This means that self-reporting of intelligibility can be highly subjective, as it is partly dependent on collective language attitudes. As a result, self-report may result in a distorted appreciation of

¹The Suriname creoles are notoriously different from the majority of the others.

²Instead of “native”, we prefer “dominant” language as this language or language may change over the lifetime of a person (Chernobilsky 2008: 431).

³We use the now-controversial term “informants” instead of, for example, “consultants” employed by, e.g., Bowerman (2008: 125), for the sake of convenience; and, more importantly, for covering distinct functions in the data gathering process.

⁴For the sake of clarity, we employ “EP/C” for one variety or language, “EP/Cs” when more than one is referred to, and “EP/C(s)” when applicable to one or more varieties/languages.

intelligibility and language practices (see also Section 4). Exposure to the other speech varieties leads to what can be termed ‘acquired intelligibility’, whereas ‘inherent intelligibility’ stems from a genetic relationship between distinct varieties, although this dichotomy has at times been considered problematic (Bouwer 2007: 257–258).

We now turn to PLURICENTRIC languages, a concept which we need to elucidate in order to explain why we have opted for referring to Atlantic EP/Cs as PLURIAREAL languages (Scheuringer 1996) in the (re)making. Kloss (e.g., 1978: 66) introduced the term “pluricentric” for standard languages with several varieties with equal distribution as, for example, official and administrative languages in distinct independent countries, e.g., Portugal and Brazil, as well as other Portuguese ex-colonies.⁵ Several criteria have to be fulfilled for a language to be considered pluricentric. These are, essentially, the following:

- The language occurs in usually several but at least two nations that function as “interacting centers”.
- Official status at state (sole official or co-official language) or regional level, allowing for the setting of norms through these interacting centers.
- Linguistic distance (abstand as defined by Kloss 1967) that distinguishes it from other languages and may serve as a symbol of linguistic and cultural identity.
- Acceptance of the pluricentric nature but also the perceived relevance of the national/regional variety for specific communities.
- Codification completed or in progress coupled with dissemination by institutions and model speakers, e.g., through formal education. (cf. Muhr 2012: 20; Clyne 1992: 1).

The concept of pluricentricity is, to some extent, problematic. Despite endorsing the existence of varieties, for example, within the Francophonie; for mostly historical reasons only one is seen as (genuinely) legitimate; distinctions may be political and not linguistic; variants cross national borders; and there is no “pure” variety (cf. Vigouroux 2013: 383). The criteria most problematic for Atlantic EP/C to be perceived of as a pluricentric language are official status and

⁵When compared to, e.g., English, the formation of standard varieties is still very much in progress, cf. da Silva et al. (2008), da Silva Sobrinho (2009), Gonçalves (2013).

codification, necessarily having an impact upon the acceptance of its pluricentricity as well as the creation of *abstand* through *ausbau* in the terminology established by Kloss (1967), which is frequent in standardization and codification. Indeed, “[p]luricentricity refers to the development of multiple standards, often national standards of a given language, while ‘pluri-areality’ downplays if not negates, any national level” (Dollinger 2019: 7).

It could be argued that “[p]luriareality merely says there is regional variation” (Dollinger 2019: 7), thence making it a basically useless concept for some scholars. Nevertheless, it is employed by others who wish to focus on linguistic differences in these language forms independent of national and political borders (Niehaus 2015, Elspaß et al. 2017). As in the case of the pidgin-creole divide, we have opted for “pluriareality” for what we consider relative neutrality especially *vis-à-vis* official status and codification.

Our study is based on the existence of a great number of shared structures of Atlantic EP/C(s) (Alleyne 1980, Holm 1989, Holm & Patrick 2007, Michaelis et al. 2013, Mühleisen 2018, etc.).⁶ As a matter of fact, it is possible to go beyond the Atlantic region and actually speak of worldwide features of EP/C(s). Faraclas et al. (2023), building on previous work, compile a list of 153 features. At least in the Atlantic region, these shared features bolster the observation that mutual intelligibility between EP/Cs exists. They are attributed to possible common origin as well as a physical, cultural, and an increasingly virtual diaspora (cf. Hancock 1986, Mair et al. 2015, McWhorter 1996).⁷

Finally, we are interested in whether this could eventually lead to the “making” or (re)construction of a language system with fluid borders for use in wider contexts, understood here in terms of Makoni & Pennycook (2005) on the disinvention and subsequent reconstitution of “language”. This may be a conscious or unconscious process where language awareness plays a significant role in the creation of imagined units with clear-cut boundaries, labels, names, and neatly defined norms established by distinct agents through at times overlapping mechanisms (both top-down and bottom-up; cf. Hüning & Krämer 2018). The question arises as to how language is deployed in advancing social agendas and in constructing identities in specific communities of practice and praxis. After all,

⁶Frequently cited examples of shared Atlantic EP/C features would be the equative copula *da/na/a*, the locative copula *de*, modal *fi/fo/fu/fa*, the 2PL pronoun *unu*, the ANT marker *bin*, and adverbial *se(l)f* ‘even’ (McWhorter 1995).

⁷Critique of the common origin scenario, called the “Domestic Hypothesis” by Hancock (1986) has been formulated by, for example, Baker (1999) and Huber (1999b) based on historical records.

in the case of Atlantic EP/Cs, we are potentially dealing with a community of 130,000,000 speakers (Faraclas 2020).

In the following section, we shall present the details of our methodology followed by the results of our online survey, interviews, and reading comprehension and translation task in Section 3.

2 Methodology

This paper relies on a multi-method approach, integrating online surveys and face-to-face interviews, which employs self-reported data, combined with a more objective, experimental reading comprehension and translation task. While self-reporting is a great tool to collect individual assumed beliefs and attitudes relatively quickly, it has its limitations. Responses may be biased based on one's perceptions, and subject to societal expectations. To mitigate the aforementioned limitations, the experimental comprehension and translation task was employed to test actual linguistic competences. Nonetheless, the paper is largely based on descriptions of language attitudes, perceptions, and intelligibility, rather than on grammatical analyses accounting for the intelligibility or lack thereof. Details of the three methods are provided in Sections 2.1–2.3.

2.1 The online survey

Our online survey was conducted from June 15, 2019 until October 9, 2019, via the platform purdue.ca1.qualtrics.com.⁸ A total of 56 informants were included in the analysis, with the breakdown by nationality and gender as outlined in Figure 1. Three nationalities (Ghanaian, Guyanese, and Nigerian) were represented, 24 of the informants being male and 32 female,⁹ with their ages ranging from 18 to above 50. An informant who identified as Ghanaian-British was grouped with the Ghanaians for sake of clarity and for the reason that her answers did not stand out in any way from the answers of the Ghanaians. When processing the results, informants' answers on their own variety were eliminated.

The informants were presented with audio samples of Ghanaian, Nigerian, Jamaican, and Sranan EP/C. The Ghanaian sample is from a TV show (*Things We Do for Love*) that ran in the early 2000s and the selected scene features a humorous conversation between two men, one of whom had come to borrow

⁸The survey was available at: https://purdue.ca1.qualtrics.com/form/SV_5cnFhaMpzwc2zKB, accessed November 12, 2019.

⁹The category "other" now commonly used in sociolinguistic research was excluded from this survey for the same pragmatic reasons as, for example, using the label "EP/C".

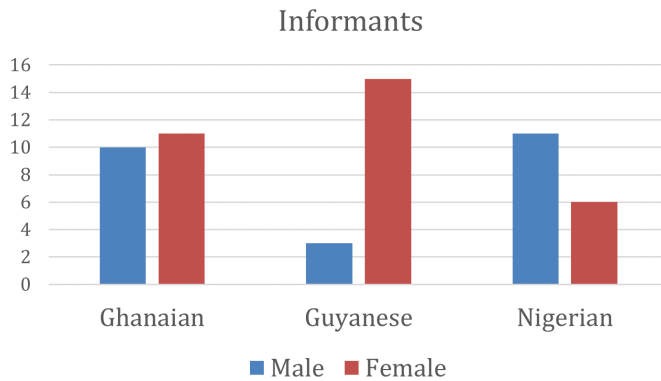


Figure 1: Breakdown of participants in the electronic survey

money and clothes in order to take a girl out on a date. The Nigerian clip is about two men fighting over a telephone recharge card after finding out the last person to recharge it had won a substantial amount of money. The Jamaican one is an interview taken from a news interview of a female Kingston resident voicing her opinion and pleading for justice after her house had been flooded the night before. Finally, the Sranan clip retells the biblical story of Noah's ark. All four samples were taken from YouTube.¹⁰

In addition to the results discussed in Sections 3.1–3.3, informants were also asked about their attitudes towards their own variety of EP/C and those of other varieties they were familiar with. However, these results are dealt with in detail in the discussion section (i.e., Section 4).

2.2 The interviews

In order to supplement the data gathered by means of the survey, 20 interviews were conducted in Guyana in October 2019, with informants representing eight different nationalities and residing in the country at the time. The breakdown by nationality and gender is outlined in Figure 2. The sample was balanced as far as gender is concerned: There were ten female and ten male informants.

The age range of the informants was from 31 years to over 61 years. The informants had lived in Guyana from less than one year up to over 10 years. In-

¹⁰Ghanaian sample: <https://www.youtube.com/watch?v=bC9I79DOmC8>, last accessed 7-28-2020; Nigerian sample: <https://www.youtube.com/watch?v=D69cixGHRIw>, last accessed 7-28-2020; Jamaican sample: <https://youtu.be/J6uwQ-HmgD8>, last accessed 7-28-2020; Sranan sample: <https://www.youtube.com/watch?v=BmVFzVUAr88>, no longer available 7-28-2020.

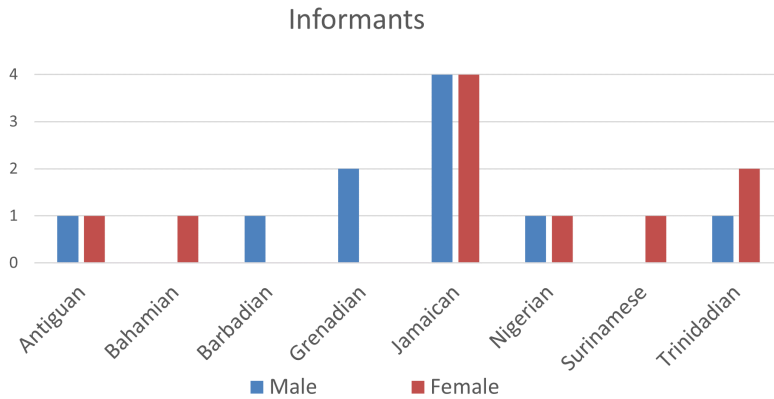


Figure 2: Breakdown of participants in the interviews

formants were selected based on convenience sampling, and were asked about their abilities to understand Guyanese Creole, specific areas of ease and difficulty, examples of differences between varieties, etc. The interviews also collected perception data on whether different varieties of EP/Cs can share a common writing system.

2.3 The reading comprehension and translation task

The reading comprehension and translation task was conducted in December 2020. Short passages were presented in eight EP/Cs: Ghanaian Pidgin, Guyanese Creole, Jamaican, Krio, Nicaraguan Creole, Nigerian Pidgin, San Andrés Creole, and Sranan. A total of 34 informants were included in the analysis, with the breakdown by nationality and gender as outlined in Figure 3. A total of 24 of the informants were Guyanese and 10 Ghanaians. All informants were university students, age ranging from 18 to 50 years.

The informants were required to identify the languages, say what the short excerpts were about and rewrite the first two lines of each excerpt in English. Additionally, informants were tasked with indicating words which they would have written differently in their language, words which were found to be difficult to understand, and rating the relative ease or difficulty in understanding the paragraph. Finally, a judgement call was made on whether they liked the writing system or not.

We now turn to the results of the distinct samples.

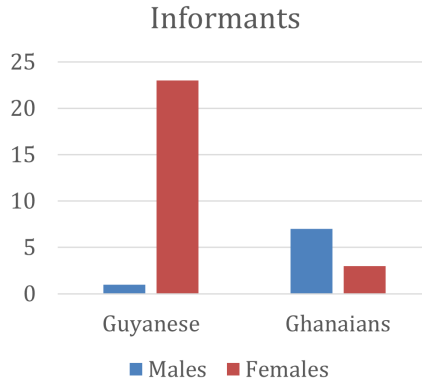


Figure 3: Breakdown of participants in the reading comprehension and translation task

3 Results

In this section, we detail the results from the online survey (Section 3.1), the interviews (Section 3.2), and the reading comprehension and translation task (Section 3.3).

3.1 The online survey

Here we provide an analysis of the data collected from the Ghanaian, Guyanese, and Nigerian participants on the Ghanaian, Nigerian, Jamaican, and Sranan samples.

3.1.1 Ghanaian sample

The breakdown of the results we obtained when we asked the informants to identify Ghanaian Pidgin can be seen in Figure 4.

In their answers, the majority of the Guyanese (11/18) picked Africa, while six selected Nigeria and one indicated that the speakers were from either Nigeria or Ghana. The majority of Nigerians (10/17) understandably picked Ghana as they are in closer contact with the variety. Of the remaining 7 Nigerian informants, 2 cited both Ghana and Nigeria; 1 could not decide between the two countries; 1 was certain that the speakers were from Nigeria; and 3 could not specify which country the Ghanaian speakers were from (with 2 selecting West Africa, while 1 simply picked Africa).

4 *Postulating Atlantic English Pidgin/Creole as a pluriareal language*

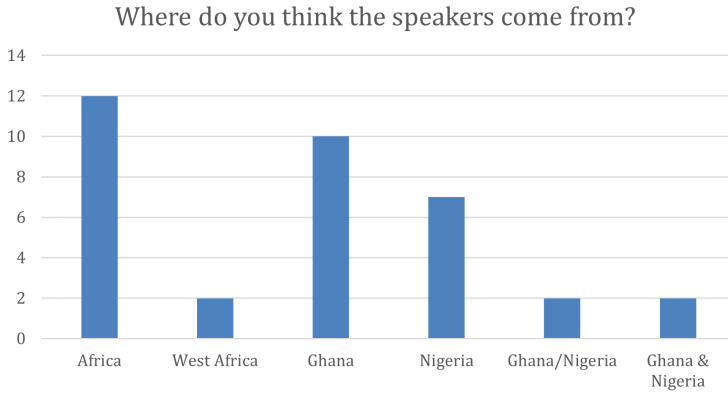


Figure 4: Variety identified by informants for the Ghanaian sample

The informants were also asked how difficult or easy it was for them to understand the speakers of the sample.

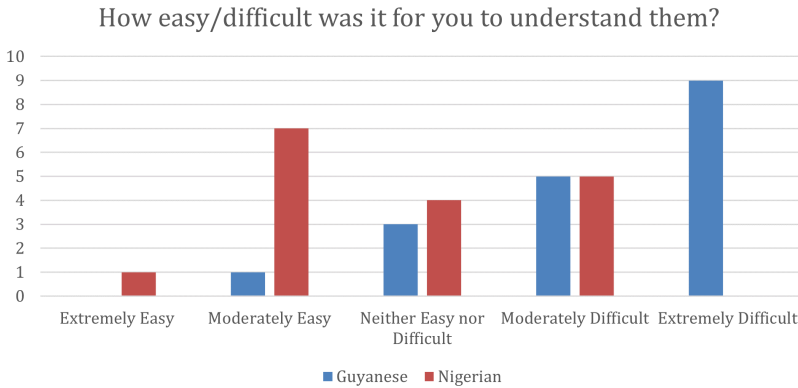


Figure 5: Difficulty of understanding Ghanaian speakers

As can be seen from Figure 5, as expected, the Guyanese informants had more difficulty in understanding the Ghanaian speakers than the Nigerians did. More precisely, the majority of the former (14/18) found the sample to be difficult (extremely or moderately so). There were also a number of Nigerian informants (6/17), including the one who said the sample was either Ghanaian or Nigerian, who found the speech moderately difficult to understand. Of the four Nigerians who found the sample neither easy nor difficult to understand, two had earlier

identified it as coming from both Ghana and Nigeria, one had said Nigeria, while one had identified it as Ghanaian.

Particular difficulties cited were constituted by what the informants identified as “indigenous words”. One such word which was frequently mentioned was *charley* (spelt variably as *chale/charlie*), which is typically one of the most noticeable vocabulary items to those who hear Ghanaian EP/C for the first time, and whose function, according to Huber (1999a: 276), is “to keep the channel of communication open”. Other vocabulary items which posed a challenge to the informants were (understandably) slang items (such as *dog-gone* ‘stupid’) and personal names (such as *Marsha*) that were thought to be content words. In addition to stating what they found difficult to understand, informants were also asked to provide alternative ways of saying (in their own EP/C) some of the things they had understood from the audio. Here also, most of the informants noted vocabulary differences between their own variety and Ghanaian EP/C; for example, one of the Nigerians said he would use *chop* ‘eat’, instead of *munch*, which was used by one of the Ghanaian speakers. One of the Guyanese informants said, instead of *mek you no mind am* ‘don’t mind him’, she would say *na bada wit he*. When discussing the Ghanaian audio, a Guyanese informant mentioned *jiggaboo* and *kangalang*, both essentially meaning ‘a stupid person’, the latter also meaning ‘hooligan, ruffian, an undisciplined person’. Nevertheless, these lexical items are not widely known to Ghanaian speakers; they would be more likely to use *dog-gone* or another slang item (for example, *john*) which means the same thing.

3.1.2 The Naija sample

As far as the Nigerian Pidgin or Naija sample is concerned, the same questions were asked.

The majority of Ghanaians (20/21) identified the variety as from Nigeria, with only one indicating that he thought the speakers were from Ghana. Again, this is expected given the geographical and cultural proximity. The majority of Guyanese (12/18) identified the variety as from somewhere in Africa but five of them also specifically cited Nigeria, something that was not the case with the Ghanaian sample. Indeed, even the single undecided Guyanese informant indicated that the speakers could be Nigerian (or Ghanaian). It can be hypothesized that this is a result of Guyanese speakers being more exposed to Nigerian than Ghanaian popular culture, especially Nollywood film productions. This is also reflected in the results presented as Figure 7. “Moderately easy to understand” was the most common answer for all groups: 14/21 of the Ghanaians and 9/18 of the Guyanese.

4 Postulating Atlantic English Pidgin/Creole as a pluriareal language

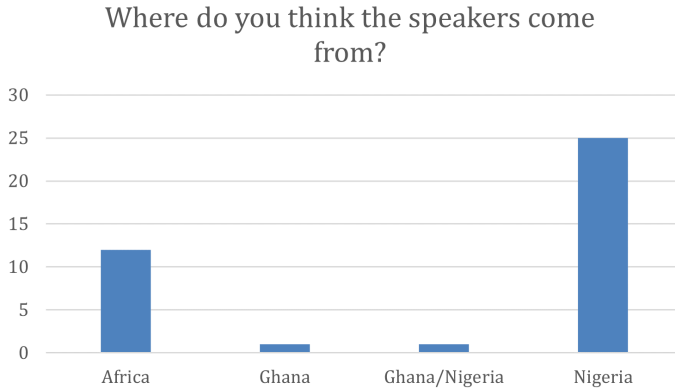


Figure 6: Variety identified by informants for the Nigerian sample

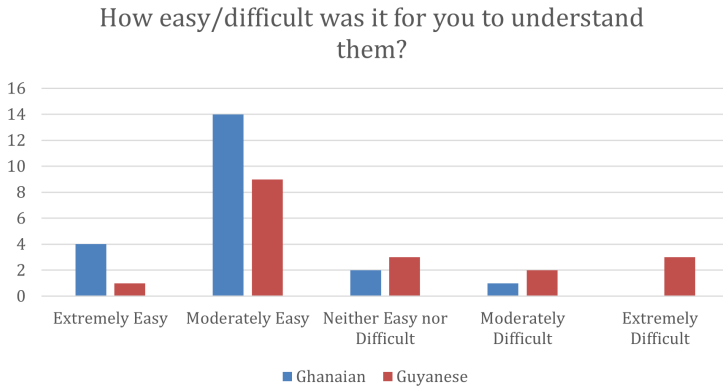


Figure 7: Difficulty of understanding Nigerian speakers

Among the special difficulties, informants cited pronunciation and rapidity of speech as well as indigenous words, e.g., *oya* ‘all right, okay’ (cf. Yoruba *oya jeka lo* ‘all right, let’s go’), which is very noticeable because of its role in discourse and frequency of occurrence. This is distinct from the Ghanaian sample where the main difficulty for informants was constituted by the lexicon. With regard to what informants would say differently in their own varieties, responses revealed (as expected) regional differences, often based on adstrate languages. For example, one of the Ghanaian informants pointed out that instead of the Nigerian Pidgin *na im* ‘he’s the one’, he would say *ibi him*. This is consistent with Huber’s (1999a: 92) observation that the equative copular *na* is noticeably absent

from Ghanaian Pidgin, whereas it is found in the other West African EP/Cs. For many Ghanaians it is indeed an emblematic marker of Naija, the variety they are most familiar with (cf. Section 4), alongside with the completive marker *don(e)*. Another example, this one from a Guyanese informant, is saying *yuh battry dead already* instead of *battery done die*, which was used by one of the Nigerian Pidgin speakers in the audio file. However, Guyanese actually also employ *don* in this context.

3.1.3 The Jamaican sample

As can be seen from Figure 8, the sample was surprisingly well identified by most informants (18/18 Guyanese, 15/21 Ghanaians and 9/18 Nigerians); with the majority of those who could not place the variety as Jamaican, still identifying it as from the Caribbean. This must be due to the fact that Jamaican speech and culture are known and imitated worldwide, albeit at times in a stylized manner, especially as a result of the spread of reggae (cf. Tomei & Hollington 2018; Moll 2017: 72), and may therefore stand for Caribbean varieties for speakers from outside of the region, e.g., West Africa (see Section 4).

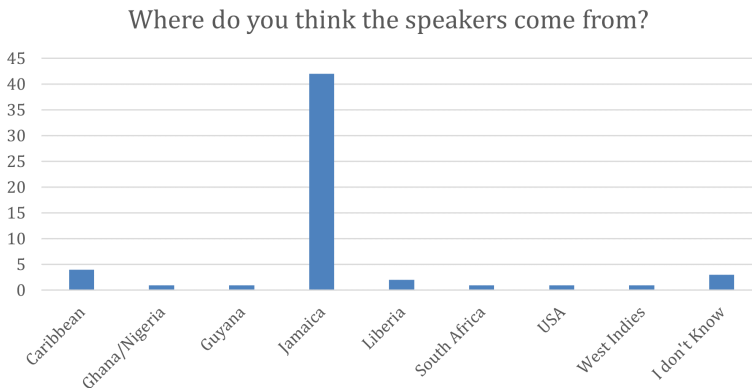


Figure 8: Variety identified by informants for the Jamaican sample

When considering the ease of understanding Jamaican, Guyanese consultants understandably scored the highest as can be gleaned from Figure 9. The majority of them (14/18) found Jamaican extremely easy to understand, while the rest found it moderately easy (3/18) or neither easy nor difficult (1/18). Globally speaking, speakers of Naija appeared to have less difficulty than Ghanaians, which may point into the direction that the influence of Jamaican (pop) culture is more

widespread in Nigeria than Ghana. More precisely, the sample was judged moderately difficult (11/21) or extremely difficult (6/21) to understand by the majority of Ghanaians; whereas a whole range of answers was given by Nigerian informants, “moderately easy” being the most frequent (6/17), followed by “moderately difficult” (5/17).

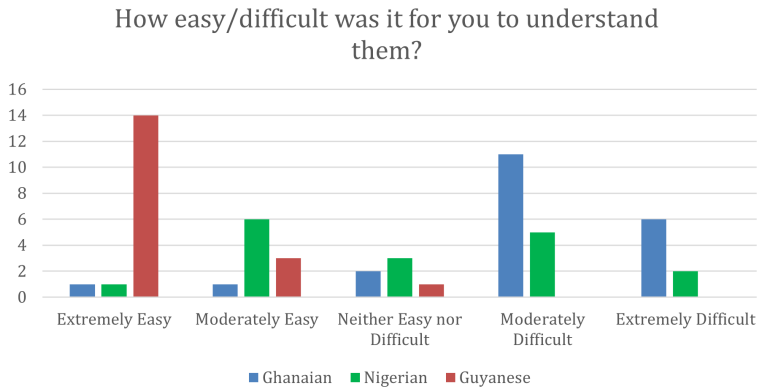


Figure 9: Difficulty of understanding Jamaican speakers

Among the difficulties mentioned in this case, we find again vocabulary as well as pronunciation, e.g., *wombo* < *waahn bak* ‘want back’, and phrasal expressions, e.g., *kom ina di yaad* ‘enter/visit the yard’. Indeed, the majority of those who found the sample difficult to understand stated that they were unable to write down specific words or phrases (possibly because they could not isolate them from the stream of words; this also applies to both examples given above). Regarding what informants would say differently in their variety, the main difference appeared to be vocabulary items such as one of the Guyanese informants pointing out that “they say *ina* and we say *in*” and a Nigerian saying he would use *for area* instead of *ina di yaad*.

3.1.4 The Sranan sample

As suggested in Section 1, Sranan (also called Sranantongo) was likely to be the EP/C most difficult to identify and understand. This hypothesis is borne out by the results we present in Figures 10 and 11.

Clearly, the variety was very difficult to identify – as witnessed by answers such as “India” (three Guyanese, one Ghanaian, and one Nigerian informant), “Some Island”, “Haiti”, etc. The category “others” includes the following responses:

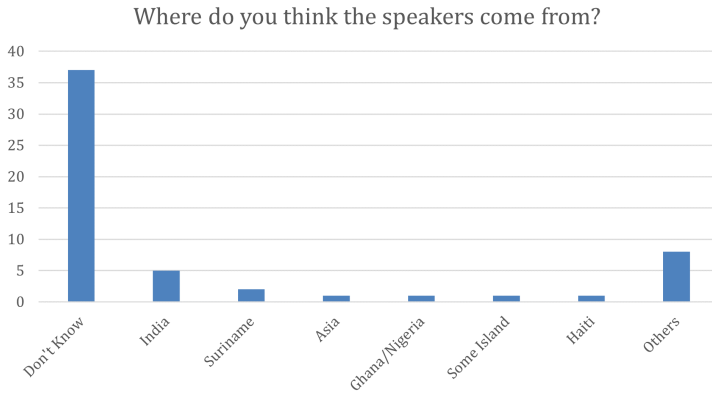


Figure 10: Variety identified by informants for the Sranan sample

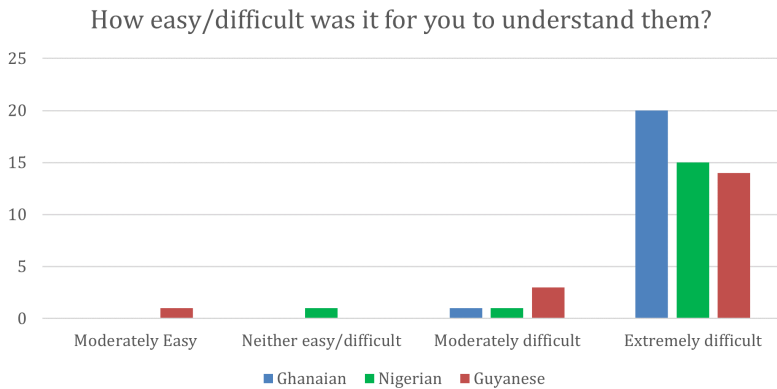


Figure 11: Difficulty of understanding Sranan speakers

“Cape Verde”; “China”; “Guyana”; “Liberia”; “Papua New Guinea”; “Some part of Mexico”; “South Africa”. Suriname was mentioned by two Guyanese who must be familiar with Sranan due to geographic, cultural, or historical proximity. As far as the difficulty of understanding Sranan is concerned, we find the following scenario where only one Guyanese, who was among the two previously cited informants, found the sample moderately easy to understand.

When asked about specific difficulties, one (Ghanaian) informant mentioned the lexical item *boato* ‘boat’, but the rest of the informants typically indicated that everything that was said was difficult.

3.2 The interviews

Results from the interviews revealed that a total of 11 informants said they understood Guyanese Creole (also called Creolese) at arrival and 18 at the time of conducting the interviews. The time range mentioned by those who have acquired passive knowledge (or receptive competence, in the terminology of Romaine 1994: 23) of Guyanese Creole was cited as from two months to three years. Nevertheless, the question on the length of acquisition was open and answered subjectively (see also Section 1). For example, a Nigerian informant stated he had acquired passive knowledge in two months after arriving in Guyana, while an informant from Suriname, despite being in Guyana for over 10 years, stated that she still did not understand Guyanese Creole. This appreciation of one's receptive competence ("very difficult to understand"; see Section 3.2.1) was shared by another female informant from Jamaica who had also resided in the country for over ten years. Could this be linked to the finding that in western-oriented societies, female speakers tend to be more conscious of linguistic norms and differences in register (cf. Labov 1972: 243; 2001: 266; Lakoff 1973: 48)?¹¹

3.2.1 The ease/difficulty of understanding Guyanese

Figure 12 depicts the relative difficulty experienced by the interviewees in understanding Guyanese Creole on a scale ranging from 1 (very easy) to 5 (very difficult). One informant chose not to comment on this matter.

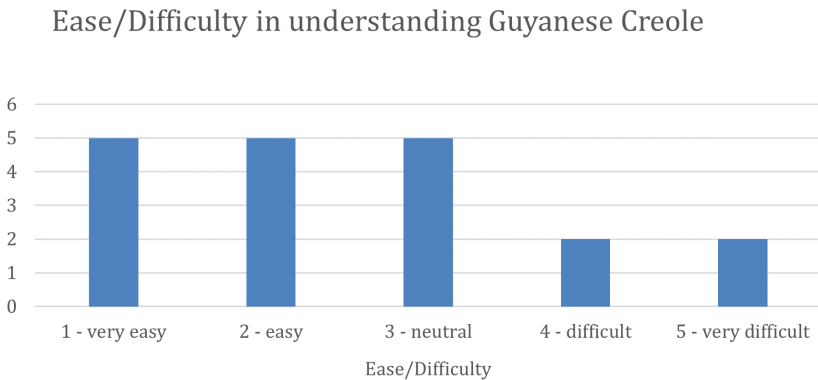


Figure 12: Relative ease/difficulty of understanding Guyanese Creole

¹¹See Keenan (1989) on Malagasy speakers for a counterexample of the Western Gender Paradox (Labov 2001: 292–293).

It can be seen that 50% of the informants (10/20) reported that Guyanese Creole was easy or very easy to understand, while five informants were neutral, finding it neither easy nor difficult to understand Guyanese Creole. It may be of interest to note that both informants who found Guyanese Creole very difficult to understand are females who have been living in Guyana for over 10 years. The difficulties mentioned included specific vocabulary items and pronunciation. Among the former, for example the following were cited: *ai paas* 'disrespect', *gyaaf* 'chat', *jos nou* 'soon', *bok maan* 'Amerindian', *balaanjii* 'eggplant', *oplif* 'collect', *push yo bodii* 'persevere', *yuuzin* (at restaurant) 'having/eating in', *tek out* (a photo) 'take', and a range of names for various fruits and vegetables. For pronunciation, informants mentioned lengthened vowels as in *maan* 'man', insertions as in *giyorl* 'girl', elision as in *wam* 'what's happening', and r-fullness as in *korna* 'corner'.¹² As in the case of the Naija sample mentioned in 3.1.2 above, rapidity of speech was also mentioned as a factor making it relatively difficult to understand Guyanese.

On the other hand, there were also linguistic features that were identified as helpful in understanding Guyanese, particularly words and structures similar to other EP/Cs, such as *gyal* to mean 'girl', the use of *dem* as a plural marker, invariability of the 3rd person object and subject pronouns, and general English-based lexical items. Additionally, the informants tend to rely on the context of utterance to decipher seemingly unfamiliar terms or structures. It was also reported by most informants (13/20) that speakers of Guyanese Creole would understand them when they use their variety of EP/C.

3.2.2 Language attitudes manifest in the interviews

Language attitudes can be divided into three components: 1. cognitive or knowledge; 2. affective or evaluative; and 3. conative or action (Agheysi & Fishman 1970: 139). Our online survey and the issues raised in the interviews discussed in the previous section essentially deal with the first component. The third is not for us as authors but eventually for the speakers to decide; we nonetheless hope to make some recommendations. In the following part of the analysis of the interviews conducted in Guyana, we deal with the affective or evaluative aspect expressed by the interviewees. To a certain extent, this was also included in the survey and the reading comprehension and translation task (see Section 4).

¹² Among the Caribbean EP/Cs, rhoticity is typical of Bajan but reported by Wells (1982: 570) for Guyanese as well.

4 *Postulating Atlantic English Pidgin/Creole as a pluriareal language*

A total of 15 out of the 20 informants reported that they like Guyanese Creole, whereas 5 reported a neutral attitude towards it. Among the answers, we find the following observations:

1. It is nice to understand other languages and cultures.
2. It is not very different from my language.
3. It is funny, sounds nice, interesting, unique, different.
4. It is not bad.
5. It is not a matter of like or dislike, I accept it.

When asked whether they liked their own variety of Atlantic EP/C, 18 answered affirmatively whereas two did not respond. Arguments in favor of one's own variety were:

1. It creates identity, unity.
2. It is great for communication.
3. It is what I grew up with, and it preserves my culture.
4. It is expressive, versatile, colorful, sentimental.
5. It makes me unique.

We were also particularly interested in the feasibility of a common writing system for Atlantic EP/C as a manifestation of its pluriareal character and part of (future) conative language attitudes. A total of 13 informants were favorable to the idea, 7 were not. Arguments cited in favor of a common writing system were that it would be practical for sharing literature and music, the linguistic background is similar, it could contribute to community cohesion (probably in the sense of the 'imagined communities' proposed by Anderson 1983), and the fact that writing is already similar (the specific example of tweets was cited, something we will be looking at in the future). The informants not favorable to a common writing system stated it would be a challenging task to complete because of differences between distinct varieties of Atlantic EP/C and expressed their fear of losing cultural diversity and identity.¹³ It can be concluded that identity-related factors were cited both in favor and against the proposal of a common writing system.

¹³Note that this fear exists even in communities deemed to speak one creole language such as Nicaraguan EP/C (cf. Freeland 2004).

3.3 Reading comprehension and translation task

In this section we present the results of the Jamaican extract as a sample of the findings from the eight EP/Cs as gathered from our Guyanese and Ghanaian informants. A more detailed analysis is presented in another study (Bartens et al. In preparation). Jamaican was chosen to exemplify the results of the reading comprehension task because of its worldwide diffusion (see Sections 3.1.3, 4, & 5). In addition, Jamaican is of particular interest for our purposes as the writing system is fairly phonemic (see, for instance, the examples below) and quite well established (see Cassidy & Le Page 1967).

As detailed in Figure 13, 90% of the Guyanese informants were able to correctly identify the Jamaican text. This proved to be a bit more difficult for the Ghanaians as only 50% were able to correctly identify the language.

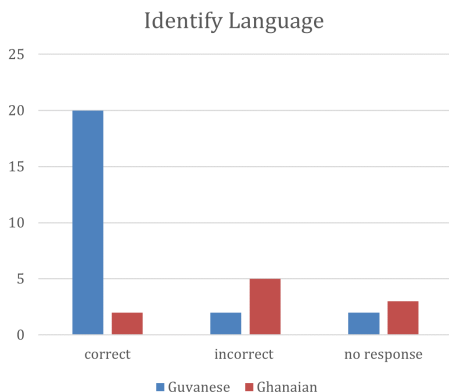


Figure 13: Identification of the Jamaican excerpt

While the Guyanese demonstrated that their understanding and translations (67% and 50%, respectively) were highly acceptable or acceptable, for the Ghanaians, the results revealed the contrary. As shown in Figure 14, 60% of the Ghanaians did not respond to these two questions, and of the other 40%, the response from 30% of the informants was not acceptable. This shows that the Ghanaians did not understand the Jamaican excerpt.

Half of the Ghanaians, as detailed in Figure 15, reported that the Jamaican data was very difficult to comprehend, while 63% of the Guyanese were neutral or indicated that it was easy to understand.

As a result, it is not surprising to note (Figure 16) that 70% of the Ghanaians, as compared to only 33% of the Guyanese, reported not liking the writing system.

4 Postulating Atlantic English Pidgin/Creole as a pluriareal language

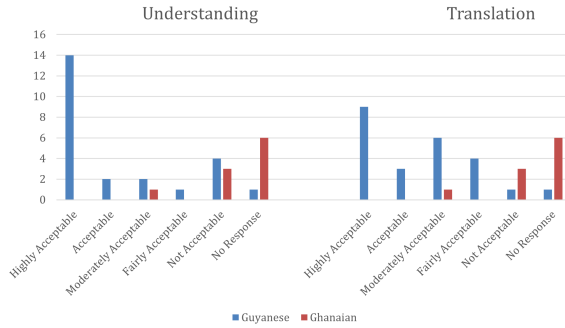


Figure 14: Understanding and translating the Jamaican excerpt

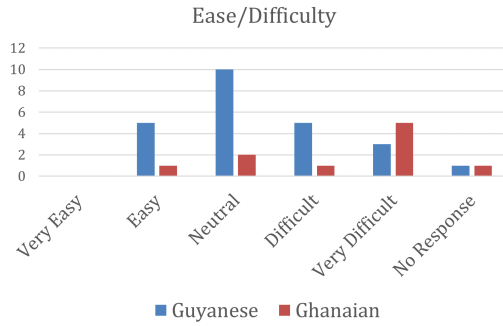


Figure 15: Difficulty of understanding the Jamaican excerpt

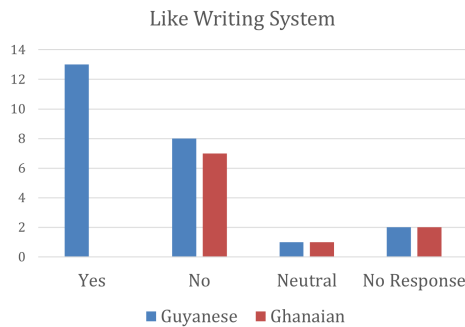


Figure 16: Liking the writing system used in the Jamaican excerpt

Examples of Jamaican words which are different or are written differently across languages were the third person singular pronoun *im* in Jamaican which is reflected as *ii* in Guyanese Creole; the lexical verb ‘eat’ which is *nyam* in Jamaican, but Guyanese informants suggested *iit*, and also ‘little’ which is *likl* in Jamaican but *lil* and *little* in Guyanese and Ghanaian, respectively. The Ghanaians also reported the English spelling for ‘tree’ and ‘try’ rather than the Jamaican *chrii* and *chrai*. The words *eleva* ‘hell of a’, *guda* ‘probably’, and *kies* ‘case’ are examples of Jamaican words that were difficult for the Guyanese informants, while *nyam* ‘eat’, *Alis* ‘Alice (name of someone)’ and *chrech* ‘stretch’ were identified as difficult for the Ghanaians.

4 Discussion

Based on the literature on Atlantic EP/Cs, it is no surprise that Sranan was found to be the most difficult variety to identify and understand. Familiarity with a variety, either as a result of geographic proximity or cultural influence, played a major role in the evaluations made by our informants. Here the online survey results stand out as it is more difficult to estimate the real contact informants have with a given EP/C as opposed to the interviews: self-report, especially in the form of written/online surveys, can notoriously distort the big picture (cf. Codó 2008: 171; see also Section 1).¹⁴ The findings from these were therefore checked against the more objective reading comprehension and translation task. The accuracy of the results gained from the self-reports is validated, as far as possible, by the actual competency demonstrated in the comprehension and translation task. The fact that Sranan was difficult to identify and understand apparently independent of the data gathering method consolidates our hypothesis of its linguistic distance from the other EP/Cs studied here.

Geographic and cultural proximity explains mutual identification and perceived intelligibility between Ghanaian and Nigerian Pidgin as well as the ease with which the Guyanese consultants identified and understood Jamaican. In these cases, we are talking about acquired intelligibility as discussed in Section 1, whereas the postulate of inherent intelligibility requires the existence of a genetic relationship as suggested, for example, by Hancock (1986) and McWhorter (1996). In fact, in the Atlantic EP/Cs, structural similarity crosses with cultural proximity

¹⁴Survey answers also quite typically feature inconsistencies, which would be detected in, for instance, ethnographically oriented fieldwork. For example, one Ghanaian informant initially did not list “Broken English” (i.e., Ghanaian EP/C) as one of the languages he speaks but stated in subsequent affirmations that he actually speaks and likes it.

and sociolinguistic accessibility. At the same time, structural similarity facilitates cognitive recognition. Therefore, the measuring of the cognitive component of language attitudes as postulated by Agheyisi & Fishman (1970: 139) is relevant for our study.

As stated above, the cultural presence of Naija, e.g., through Nollywood productions, in the Caribbean probably explains why the Guyanese informants were more familiar with it than with Ghanaian Pidgin. Based on this sample, Jamaican also appears to be culturally more present in Nigeria than in Ghana, and may be additionally identified by African informants as the Caribbean EP/C by default (cf. Section 3.1.3).

While the previous discussion is essentially based on the online survey, the interviews conducted in Guyana confirmed that time of exposure is a key factor: Most informants found Guyanese Creole easy to understand within less than three years. What constitutes passive knowledge is again subjective and based on the standards and goals an individual sets for themselves in an intercultural context (cf. Dörnyei & Ottó 1998, Dörnyei & Csizér 2005). At the same time, most interviewees not only displayed positive attitudes towards both Guyanese Creole and their own variety but also supported the idea of a common writing system. This is consistent with the views of the online survey informants as the majority of them (44 out of the 47 who responded to that question) reported positive attitudes towards their own varieties. Additionally, out of the 51 online informants who answered the question on the idea of a common writing system, the overwhelming majority (i.e., 48) were in favor of such a development.

Arguments in favor of and against a common writing system appear to concentrate on identity-related issues. The major difficulties perceived were not only in the field of vocabulary, typical of any pluriareal or pluricentric language,¹⁵ but also in pronunciation. Pronunciation is easily mistaken for – but also closely related to – orthographies and writing systems, especially if these are devised for languages with a short tradition of writing. Schieffelin & Doucet (1994: 192–193) state that:

When a language is codified and an orthography is officially adopted, this is usually interpreted to mean that there is one correct way to spell and write the language, and that all others are simply wrong. [...] And when a variety through its officialization is given the status of a standard, the users of the other varieties sometimes react with surprising virulence because they feel that their language variety and its speakers are denied representation.

¹⁵In fact, Berschin et al. (2012: 382) define Spanish, a case in point, as a pluricentric language based on the lexicon, not “grammar”, i.e. language structure.

Opting for “normalization” in the sense of allowing for regional variation in the initial stages of writing as proposed by Koskinen (2010) for Nicaraguan Creole constitutes a sound option – when we are dealing with a geographically limited area.¹⁶ Our interest lies in the feasibility of writing Atlantic EP/C on a much larger scale which, as some informants’ responses reveal, may turn out a much more complicated endeavor. In this context, it is relevant to bear in mind that certain graphemes, for example {k}, {w}, and {y}, have led to heated debates (Bartens et al. In preparation) and that, on the other hand, not writing does not constitute an option in today’s world. Writing is a fundamental condition for languagehood according to present language ideologies (e.g., Lüpke 2018). In the long run, writing is not only a key element for constituting “language” in the sense of H vs. L languages/varieties (Ferraz 1979, Fishman 1967), but also for passing from a pluriareal to a pluricentric language through distinct processes of codification as outlined in Section 1. Writing may, however, be achieved not only through flexibility (see Koskinen’s interpretation of the term “normalization” above) but also through grassroots literacy (Blommaert 2008). These are possible paths of action to counter the fears of those who maintain that the inherent heterogeneity and creativity of creoles has to be sacrificed when, for instance, creating an orthography (cf. Freeland 2004: 124).

Assessing the informants’ competencies in applying different orthographies in the reading comprehension and translation task is therefore fitting to the current study. Again, geographic and cultural proximity accounts for the actual intelligibility as demonstrated within the Caribbean and West African region. However, delving deeper into the intricacies of the writing systems employed for each language, we noted that informants who indicated that they liked a specific writing system liked it because they could read/understand it; the sole exception here was Krio. While 75% of the Guyanese and 40% of the Ghanaians demonstrated that they understand Krio, only 25% and 10% Guyanese and Ghanaians respectively indicated that they liked the writing system. This is because while Krio uses the Latin script which the informants are accustomed to, it also employs three phonetic letters – {ɛ}, {ɔ}, and {ŋ} – which were reported to be difficult.

The orthography used to write Sranan was not liked by any of the informants, though it was understood by 21% of the Guyanese informants. The reason for not liking the writing system of Sranan was simply because the language was seen as difficult and the informants did not understand it. This is so because most of the words in Sranan, though English-based, do not look like their English etymons

¹⁶DeGraff (2014: 313) cautions that “uniform conventions for written representations [...] [should] not distract from other valuable priorities that function to promote the language”.

and structural and grammatical contrasts may obstruct intelligibility. Sranan has been in contact with Dutch rather than English and therefore, unlike other EP/Cs, it has no post-creole continuum and did not get to decreolize. (See Sebba 2000 for a full description of Sranan's orthography.)

The Guyanese informants were more likely to report that they liked the Guyanese, Nicaraguan, and Jamaican writing systems (83%, 75%, and 54%, respectively) when compared to the other languages. Conversely, Ghanaian informants were more likely to report that they liked the Nigerian and Ghanaian writing systems (70% and 50%, respectively). The structural and lexical similarity of the languages is assumed to account for the cognitive recognition, comprehension, and translation of the languages in the reading comprehension and translation task. This is also very much conditioned by the familiarity with distinct writing systems. (See Winer 1990 for four different models for writing an EP/C.)

A postcolonial approach is called for precisely when we consider the necessity of writing. What is meant by postcolonial in this context? Linguists, especially, tend to associate "postcolonial" with literary studies. However, a field of postcolonial linguistics is starting to emerge and the research questions we are interested in – Atlantic EP/C as a pluriareal language, mutual intelligibility, and the possibility of writing – would fall into the domain of postcolonial sociolinguistics (cf. Makoni 2011, Levisen & Sippola 2019). For our purposes, it seems more useful to define "postcolonial" as a time-defining concept implying change (cf. Calabrese 2015: 1; Anchimbe 2018: xiii) than within the framework of power which postcolonial studies in general operate with (a position advocated, e.g., by Warnke 2017). Both stances, the descriptive-causal and the critical-reflective one, share their opposition to Eurocentrism (Levisen & Sippola 2019: 2).

Though space will not permit us to fully explore the postcolonial approach in this present paper, it is our aim to return to this in subsequent works. Nonetheless, in our understanding of postcolonial, i.e., as an essentially time-defining concept, it is important to consider language ecologies, both past (cf. The Founder Principle) and present, which draw on the feature pool available through population groups at a given time and the possibility of translanguaging in the ongoing formation of Atlantic EP/C(s) (Haugen 1971, Mufwene 1996, 2013, Schneider 2007: 22–23, García & Li 2014). For language ecologies to emerge, the concept of "community" is highly relevant. Community is a highly elusive concept (Mühleisen 2017: xii) and represents "a relation constantly under negotiation" (Brydon & Coleman 2008: 2). We envision that speakers of various Atlantic EP/Cs may actually at some point conceive of themselves as speaking a single Atlantic EP/C (i.e., a pluriareal language, albeit with varieties; see Section 1). But a (speech) community does not exist without the individuals who make it up. Indeed, Mufwene

(2013: 323) stresses the role of the individual in the ecology of language and affirms that:

[...] factoring in the speaker as the most direct external ecological factor to language, as he/she contributes variation to the emergent, ever-evolving language and participates in: 1) the spread or elimination of variants through the selections he/she makes from among the competing variants (be they languages or linguistic features), 2) the emergence of new norms, and 3) sometimes the emergence of new varieties (Mufwene 2013: 311).

We are arguing that competition and selection of variants is not necessarily one of the first steps, at least not on a larger scale, in the (re)constitution of an Atlantic EP/C which will undoubtedly retain variation as a pluriareal language. On the other hand, the emergence of new linguistic norms, including norms of usage, and their consolidation as an acknowledged diasystem, albeit with fluid borders, is a more realistic outcome. This diasystem could theoretically at some point be on its way to a pluricentric language with speakers increasingly accommodating their linguistic production in the sense of convergence towards the other speakers' varieties (cf., for example, Giles & Ogay 2007: 295 for the Communication Accomodation Theory) due to the fulfillment of the criteria stipulated, e.g., by Muhr (2012: 20) and Clyne (1992: 1; see Section 1), but for the time being this may seem to be a relatively farfetched scenario. Indeed, if we prefer to consider languages less as "linguistically defined objects" and rather as speech and resources, "the real bits and chunks of language that make up a repertoire, and [...] real ways of using this repertoire in communication" (Blommaert 2010: 5, 173), continuing to conceive of Atlantic EP/Cs as pluriareal languages may constitute a more feasible solution even in the long run. However, and as stated above (Section 1), in the "making of languages" – which may actually consist of the disinvention and subsequent reconstitution of linguistic varieties (Makoni & Pennycook 2005) – clear-cut boundaries, labels, names, and norms are required especially in cases of long-standing stigmatization and minoritization to which Atlantic EP/Cs also belong (cf. Hüning & Krämer 2018). This is especially true in sociolinguistic terms, whereas structural fluidity is necessary for use in wider contexts, including writing, when dealing with as complex a diasystem as Atlantic EP/Cs.¹⁷

¹⁷Note, however, that fluidity (and hybridity) is problematic in that, as advocated in postcolonial studies, this concept overlaps with neoliberalism (Kubota 2016). In the case of Atlantic EP/C, structural fluidity is a necessity which does not clash with the time-defining, descriptive causal perspective adopted here.

The impact and the peculiarities of virtual communities are no longer to be underestimated. Whereas cultural diffusion (cf. the above-mentioned example of Jamaican reggae) and diaspora formation lead to the spread, acceptance, and consolidation of certain varieties and their norms of usage (see Mühleisen & Schröder 2017 on the shift in role and prestige of Caribbean Creoles and Cameroonian EP/C in new diasporic metropolitan environments), new technology has opened totally new opportunities for language diffusion and development, including the democratization of the latter process (cf. Eisenlohr 2004, Bartens 2019). For example, cyberspace has developed a sociolinguistic order of its own. Mair (2019) on Jamaican Standard English vs. Jamaican Creole and Nigerian Standard English vs. Naija found that what is stigmatized on the ground often actually becomes prestigious in cyberspace. At the same time, it becomes increasingly difficult to apply the concept of “speech community” to, for example, (diasporic) web-forum interaction (Moll 2017: 91). In the case of Atlantic EP/C(s), this may actually constitute an opportunity: If instead of the classical speech community (Gumperz 1968; Labov 1972: 120–121) we focus on communities of practice in the sense of “an agreement of people who come together around mutual engagement in an endeavour” (Eckert & McConnell-Ginet 1992: 464), less competent or so-called peripheral speakers (Labov 1972) may be integrated by means of computer mediated communication.

We maintain that the affirmation “the language of the imagined global community is clearly English and the written script is digital” (Mühleisen 2017: xvii) is not true, especially not in the case of very widely spoken pluriareal Atlantic EP/C (see Section 1). Its different varieties are already widely used in digital communication and we believe that in the future this use will further increase and may lead to a consciousness of belonging to an imagined community, which is already in place among speakers of Caribbean varieties. The linguistic practices of this imagined community are able to counter elite closure by enabling alternative routes in the flow of information (cf. DeGraff 2014: 284, 311).

5 Conclusions

The point of departure for this study was the observation that many speakers of closely related varieties of Atlantic EP/Cs can understand each other. First, we defined and justified the use of certain key concepts such as Atlantic EP/C(s). Our argument is that distinct varieties, i.e., EP/Cs, could be conceived of – and in part already are viewed as – the Atlantic EP/C diasystem. As we insist, this language system has fluid borders, allowing for use in distinct contexts and domains. Throughout the paper, we were concerned with intelligibility. We also

discussed why we consider Atlantic EP/C a pluriareal language in the making and not a pluricentric one despite the fact that the former term is contested by some scholars (e.g., Dollinger 2016: 7). This (re)constitution of a language system with as many as 130,000,000 speakers is facilitated by a great number of shared linguistic structures.

Three lines of inquiry were pursued in order to verify – or falsify – the hypothesis of mutual intelligibility: 1. an online survey available from June 15, 2019, until October 9, 2019 which included audio clips of Ghanaian, Nigerian, Jamaican, and Sranan EP/C; 2. interviews of other Atlantic EP/C speakers residing in Guyana at the time of the study (October 2019); and 3. a reading comprehension and translation task conducted with university students in Guyana and Ghana (December 2020). The third line of inquiry was adopted in order to mitigate the caveats of self-report of intelligibility arising from the dependence on language attitudes (see Sections 1 and 2).

The data reveal that geographic and cultural proximity is taken to explain acquired intelligibility between Ghanaian Pidgin and Naija as well as the ease with which the Guyanese informants identified and understood Jamaican. Familiarity with Nollywood and Jamaican culture, especially reggae, were suggested to account for those varieties being better known and probably taken as representative of West Africa and the Caribbean, respectively.

Structural similarity was assumed to facilitate cognitive recognition, as was demonstrated in the reading comprehension and translation task. In the case of reading comprehension, cognitive recognition is also very much conditioned by the familiarity with distinct writing systems (see Winer 1990 for four different models for writing an EP/C). Writing was argued to constitute one of the major criteria for languagehood and therefore it is relevant to state that attitudes not only towards one's own but also other varieties as well as writing them were fairly positive. For example, 13 out of the 20 interviewees were favorable towards a common orthography for writing EP/C languages.

The question of writing can be framed in postcolonial thought. For our purposes, defining “postcolonial” as a time-defining concept seemed the most adequate approach. Despite this descriptive-causal (as opposed to critical-reflective) interpretation, we believe that it is precisely new technologies, offering new possibilities for the revitalization and development of minorized languages, which will enable Atlantic EP/Cs to consolidate into a community of practice in its own right. For the time being, we leave exploring the topic from a postcolonial angle – as well as a more in-depth discussion of the reading comprehension and translation task – for future studies.

Acknowledgements

We would like to gratefully acknowledge the invaluable contribution of Dr. Ucheonna Oyali, University of Abuja, during the initial phase of the study (online survey). We would also like to thank the two anonymous reviewers for their comments and suggestions.

References

- Agheysi, Rebecca & Joshua A. Fishman. 1970. Language attitude studies: A brief survey of methodological approaches. *Anthropological Linguistics* 12(5). 137–157.
- Alleyne, Mervyn Coleridge. 1980. *Comparative Afro-American*. Ann Arbor: Karoma.
- Anchimbe, Eric A. 2018. *Offers and offer refusals: A postcolonial pragmatics perspective on world Englishes*. Amsterdam: John Benjamins.
- Anderson, Benedict. 1983. *Imagined communities: Reflections on the origin and spread of nationalism*. London: Verso.
- Baker, Philip. 1999. Investigating the origin and diffusion of shared features among the Atlantic English Creoles. In Philip Baker & Adrienne Bruyn (eds.), *St Kitts and the Atlantic Creoles: The texts of Samuel Augustus Mathews in perspective*, 315–364. London: University of Westminster Press.
- Bakker, Peter. 2008. Pidgins versus creoles and pidgincreoles. In Silvia Kouwenberg & John Singler (eds.), *Handbook of Pidgin and Creole studies*, 130–157. Oxford: Blackwell-Wiley.
- Bakker, Peter & Yaron Matras. 2013. Introduction. In Peter Bakker & Yaron Matras (eds.), *Contact languages: A comprehensive guide (Language Contact and Bilingualism)*, vol. 6, 1–14. Boston/Berlin: Mouton de Gruyter.
- Bartens, Angela. 2019. Language ideologies in the San Andrés (and Old Providence) Raizal community on the way to democracy. *Neuphilologische Mitteilungen* CXIX(II). 393–418.
- Bartens, Angela, Tamirand Nnena De Lisser & Kwaku Owusu Afriyie Osei-Tutu. In preparation. *Writing and reading a pluriareal language: Findings from a reading comprehension and translation task in Atlantic English Pidgin/Creole*.
- Berschin, Helmut, Julio Fernández-Sevilla & Josef Felizberger. 2012. *Die spanische Sprache: Verbreitung – Geschichte – Struktur*. Hildesheim: Georg Olms/Wissenschaftliche Buchgesellschaft.
- Blommaert, Jan. 2008. *Grassroots literacy: Writing, identity and voice in Central Africa*. London: Routledge.

- Blommaert, Jan. 2010. *The sociolinguistics of globalization*. Cambridge: Cambridge University Press.
- Bouwer, Leoni E. 2007. Intercomprehension and mutual intelligibility among southern Malagasy languages. *Language Matters* 38(2). 79–98.
- Bowern, Claire. 2008. *Linguistic fieldwork: A practical guide*. Basingstoke: Palgrave Macmillan.
- Brydon, Diana & William Donald Coleman. 2008. Globalization, autonomy, and community. In Diana Brydon & William Donald (eds.), *Renegotiating community: Interdisciplinary perspectives, global contexts*, 1–28. Vancouver: University of British Columbia.
- Calabrese, Rita. 2015. Introduction: Post-colonial convergence and divergence. In Rita Calabrese, John Kenneth Chambers & Gerhard Leitner (eds.), *Variation and change in postcolonial contexts*, 1–14. Cambridge: Cambridge Scholars Publishing.
- Cassidy, Frederic Gomes & Robert Brock Le Page. 1967. *Dictionary of Jamaican English*. London & New York: CUP.
- Chernobilsky, Ellina. 2008. Language dominance. In Josué M González (ed.), *Encyclopedia of bilingual education*, 430–435. Los Angeles: Sage.
- Clyne, Michael. 1992. Pluricentric languages: Introduction. In Michael Clyne (ed.), *Pluricentric languages: Differing norms in different nations*, 1–10. Berlin: De Gruyter.
- Codó, Eva. 2008. Interviews and questionnaires. In Li Wei & Melissa G Moyer (eds.), *The Blackwell guide to research methods in bilingualism and multilingualism*, 158–176. Malden, MA: Blackwell.
- da Silva, Nilce, Domingos Carlos Buque, Luís Miguel Cardoso & Maria de Jesus Fonseca. 2008. Alguns apontamentos sobre o Acordo Ortográfico: Três olhares lusófonos – Brasil, Moçambique e Portugal. *Revista ACOALFAPlp: Acolhendo a Alfabetização nos Países de Língua Portuguesa* 3(5). 1–24. <http://www.acoalfaplp.net> (3 June, 2020).
- da Silva Sobrinho, José Simão. 2009. Acordo ortográfico e a unidade imaginária da língua portuguesa. *Revista do GEL* 6(2). 232–242.
- DeGraff, Michel. 2014. The ecology of language evolution in Latin America: A Haitian postscript toward a postcolonial sequel. In Salikoko Mufwene (ed.), *Iberian imperialism and language evolution in Latin America*, 247–327. Chicago: University of Chicago Press.
- Deuber, Dagmar. 2005. *Nigerian Pidgin in Lagos: Language contact, variation and change in an urban African setting*. London: Battlebridge. 272.

- Dollinger, Stefan. 2016. *On parallels, differences and distortions in the pluricentricity of English and German. Presentation at the Deutsch in Österreich (German in Austria) Conference: Deutsch in Österreich und andere plurizentrische Kontexte. Variation – Kontakt – Perzeption. Schönbrunn Castle, Vienna, Austria, 7 July 2016.* https://www.academia.edu/26828790/On_parallels_differences_and_distortions_in_the_pluricentricity_of_English_and_German_SCRIPTED_TEXT_and_SLIDES (3 April, 2020).
- Dollinger, Stefan. 2019. *The pluricentricity debate: On Austrian German and other Germanic standard varieties.* Milton Park: Routledge Focus.
- Dörnyei, Zoltán & Kata Csizér. 2005. The effects of intercultural contact and tourism on language attitudes and language learning motivation. *Journal of Language and Social Psychology* 24(4). 327–357. DOI: 10.1177/0261927x05281424.
- Dörnyei, Zoltán & Istvan Ottó. 1998. Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics* 4. 43–69.
- Eckert, Penelope & Sally McConnell-Ginet. 1992. Think practically and look locally: Language and gender as community-based practice. *Annual Review of Anthropology* 21(1). 461–488. DOI: 10.1146/annurev.an.21.100192.002333.
- Eisenlohr, Patrick. 2004. Language revitalization and new technologies: Cultures of electronic mediation and the refiguring of communities. *Annual Review of Anthropology* 33(1). 21–45. DOI: 10.1146/annurev.anthro.33.070203.143900.
- Elspaß, Stephan, Christa Dürscheid & Arne Ziegler. 2017. Zur grammatischen Pluriarealität der deutschen Gebrauchsstandards – oder: Über die Grenzen des Plurizentritätsbegriffs. *Zeitschrift für Deutsche Philologie* 136. 69–91.
- Faraclas, Nicholas. 2020. Identity politics. In Miriam Meyerhoff & Umberto Ansaldo (eds.), *Routledge handbook on Pidgin and Creole languages*, 269–285. Abingdon: Routledge.
- Faraclas, Nicholas, Micah Corum & Rhoda Arrindell. 2023. Subaltern currents and transgressive waves: Inter-oceanic agents hospitable connectedness and “creole languages”. In Sukla Chatterjee, Joanna Chojnicka, Anna-Katharina Hornidge & Kerstin Knopf (eds.), *Postcolonial oceans: Contradictions, heterogeneities, knowledges, materialities* (Anglophone Postcolonial Studies 1), 53–70. Heidelberg: Heidelberg University Publishing. DOI: 10.17885/heiup.1046.c17301.
- Ferraz, Luis I. 1979. *The creole of São Tomé.* Johannesburg: Witwatersrand University Press.
- Fishman, Joshua A. 1967. Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues* 23(2). 29–38. DOI: 10.1111/j.1540-4560.1967.tb00573.x.

- Freeland, Jane. 2004. Linguistic rights and language survival in a creole space: Dilemmas of Nicaragua's Caribbean Coast Creoles. In Jane Freeland & Donna Patrick (eds.), *Language rights and language survival: A sociolinguistic and sociocultural approach*, 103–138. Manchester: St. Jerome Publishing.
- García, Ofelia & Wei Li. 2014. *Translanguaging: Language, bilingualism and education*. Basingstoke: Palgrave Macmillan.
- Giles, Howard & Tania Ogay. 2007. Communication accommodation theory. In Bryan B Whaley & Wendy Samter (eds.), *Explaining communication: Contemporary theories and exemplars*, 293–310. Mahwah, NJ: Lawrence Erlbaum.
- Gonçalves, Perpétua. 2013. O Português em África. In Eduardo Buzaglo Paiva Raposo (ed.), *Gramática do Português i*, 157–178. Lisboa: Fundação Calouste Gulbenkian.
- Gumperz, John Joseph. 1968. The speech community: International encyclopedia of the social sciences. In Alessandro Duranti (ed.), *Linguistic anthropology: A reader*, 381–386. New York: Macmillan.
- Hall, Robert Anderson. 1962. The life cycle of pidgin languages. *Lingua* 11. 151–156.
- Hancock, Ian Francis. 1986. The domestic hypothesis, diffusion and componentiality: An account of anglophone creole origins. In Pieter Muysken & Norval Smith (eds.), *Substrata versus universals in Creole genesis*, 71–102. Amsterdam: Benjamins.
- Haugen, Einar. 1971. The ecology of language: The linguistic reporter, supplement 25.19–26. In Anwar S. Dil (ed.), *The ecology of language*, 325–339. Stanford: Stanford University Press.
- Holm, John. 1989. *Pidgin and Creole languages*. Cambridge: Cambridge University Press.
- Holm, John & Peter Patrick. 2007. *Comparative creole syntax: Parallel outlines of 18 creole grammars*. London: Battlebridge.
- Huber, Magnus. 1999a. *Ghanaian pidgin English in its West African context: A sociohistorical and structural analysis* (Varieties of English around the World G24). Amsterdam: John Benjamins.
- Huber, Magnus. 1999b. On the origin and diffusion of Atlantic English Creoles: First attestations from Krio. In Philip Baker & Adrienne Bruyn (eds.), *St Kitts and the Atlantic Creoles: The texts of Samuel Augustus Mathews in perspective*, 365–378. London: University of Westminster Press.
- Hüning, Matthias & Philipp Krämer. 2018. *Negotiating standards between Suriname and the Netherlands. International Conference for Postcolonial Linguistics*, Zurich, June 2018.

- Keenan, Elinor. 1989. Norm-makers, norm-breakers: Uses of speech by men and women in a Malagasy community. In Richard Bauman & Joel Scherzer (eds.), *Explorations in the ethnography of speaking*, 2nd edn., 125–143. Cambridge: Cambridge University Press.
- Kloss, Heinz. 1967. “Abstand languages” and “Ausbau languages”. *Anthropological Linguistics* 9(7). 29–41.
- Kloss, Heinz. 1978. *Die Entwicklung neuer germanischer Kultursprachen seit 1800*. Düsseldorf: Schwann.
- Koskinen, Arja. 2010. Kriol in Caribbean Nicaragua schools. In Bettina Migge, Isabelle Léglise & Angela Bartens (eds.), *Creoles in education: An appraisal of current programs and projects*, 133–165. Amsterdam: John Benjamins.
- Kubota, Ryuko. 2016. The multi/plural turn, postcolonial theory, and neoliberal multiculturalism: Complicities and implications for applied linguistics. *Applied Linguistics* 37(4). 474–494.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia, PA: University of Philadelphia Press.
- Labov, William. 2001. *Principles of linguistic change: Social factors*, vol. 2. Malden, MA: Blackwell.
- Lakoff, Robin T. 1973. Language and woman’s place. *Language in Society* 2(1). 45–80.
- Levisen, Carsten & Eeva Sippola. 2019. Postcolonial linguistics: The editors’ guide to a new interdiscipline. *Journal of Postcolonial Linguistics* 1. 1–15.
- Lüpke, Friederike. 2018. Escaping the tyranny of writing: West African regimes of writing as a model for multilingual literacy. In Constanze Weth & Kasper Juffermans (eds.), *The tyranny of writing: Ideologies of the written word*, 129–147. London: Bloomsbury.
- Mair, Christian. 2019. World Englishes in cyberspace. In Daniel Schreier, Marianne Hundt & Edgar W. Schneider (eds.), *The Cambridge handbook of world Englishes*, 360–383. Cambridge: Cambridge University Press.
- Mair, Christian, Susanne Mühleisen & Eva Ulrike Pirker. 2015. Selling the Caribbean: An introduction. *Zeitschrift für Anglistik und Amerikanistik* 63(2). 139–156.
- Makoni, Sinfree. 2011. Sociolinguistics, colonial and postcolonial: An integrationist perspective. *Language Sciences* 33(4). 680–688. DOI: 10.1016/j.langsci.2011.04.020.
- Makoni, Sinfree & Alastair Pennycook. 2005. Disinventing and (re)constituting languages. *Critical Inquiry in Language Studies* 2(3). 137–156. DOI: 10.1207/s15427595cils0203_1.

- McWhorter, John. 1995. Sisters under the skin. *Journal of Pidgin and Creole Languages* 10(2). 289–333. DOI: 10.1075/jpcl.10.2.04mcw.
- McWhorter, John H. 1996. It happened at Coromantin. *Journal of Pidgin and Creole Languages* 12. 1–44.
- McWhorter, John H. 2011. *Linguistic simplicity and complexity: Why do languages undress? (Language Contact and Bilingualism)*. Berlin/Boston: Mouton De Gruyter.
- Michaelis, Susanne M., Philippe Maurer, Martin Haspelmath & Magnus Huber. 2013. *The atlas and survey of pidgin and creole language structures*, vol. I: English-based and Dutch-based languages. Oxford: Oxford University Press.
- Moll, Andrea. 2017. Diasporic Cyber-Jamaican: Stylized dialect of an imagined community. In Susanne Mühleisen (ed.), *Contested communities: Communication, narration, imagination*, 69–94. Leiden: Brill.
- Mufwene, Salikoko S. 1996. The founder principle in creole genesis. *Diachronica* 13(1). 83–134. DOI: 10.1075/dia.13.1.05muf.
- Mufwene, Salikoko S. 2013. The ecology of language: Some evolutionary perspectives. In Elza Kioko Nakayama Nenoki do Couto, Davi Borges de Albuquerque & Gilberto Paulino de Araújo (eds.), *Ecolinguística e imaginário*, 302–327. Brasília: Thesaurus.
- Mühleisen, Susanne & Anne Schröder. 2017. Prestige change in contact varieties of English in urban diaspora communities. In Susanne Mühleisen (ed.), *Contested communities: Communication, narration, imagination*, 51–68. Leiden: Brill.
- Mühleisen, Susanne. 2018. Comparative pronominal systems: Cognitive and pragmatic functions of second person plural pronouns in West African Pidgins and Caribbean Creole. In Universidad Nacional de Costa Rica (ed.), *Connecting the Caribbean: Languages, borders, and identities*, vol. 5. Paper presented at SCL & SPCL 22nd Biennial Conference.
- Muhr, Rudolf. 2012. *Non-dominant varieties of pluricentric languages getting the picture*. In collaboration with Catrin Norrby, Leo Kretzenbacher, Carla Amorós Negre. In memory of Michael Clyne. Frankfurt am Main: Peter Lang.
- Mühleisen, Susanne. 2017. Introduction: On community formation, manifestation, and contestation acts of membership and exclusion. In Susanne Mühleisen (ed.), *Contested communities: Communication, narration, imagination*, ix–xxiv. C. Leiden: Brill.
- Niehaus, Konstatin. 2015. Areale Variation in der Syntax des Standarddeutschen: Ergebnisse zum Sprachgebrauch und zur Frage Plurizentrik vs. Pluriarealität. *Zeitschrift für Dialektologie und Linguistik* 82(2). 133–168.

4 *Postulating Atlantic English Pidgin/Creole as a pluriareal language*

- Romaine, Suzanne. 1994. *Language in society: An introduction to sociolinguistics*. Oxford: Oxford University Press.
- Scheuringer, Hermann. 1996. Das Deutsche als pluriareale Sprache: Ein Beitrag gegen staatlich begrenzte Horizonte in der Diskussion um die deutsche Sprache in Österreich. *Die Unterrichtspraxis / Teaching German* 29(2). 147–153.
- Schieffelin, Bambi B & Rachele Charlier Doucet. 1994. The “real” Haitian Creole: Ideology, metalinguistics, and orthographic choice. *American Ethnologist* 21(1). 176–200. DOI: 10.1525/ae.1994.21.1.02a00090.
- Schneider, Edgar W. 2007. *Postcolonial English: Varieties around the world*. Cambridge: Cambridge University Press.
- Sebba, Mark. 2000. Orthography and ideology: Issues in Sranan spelling. *Linguistics* 38(5). 925–948. DOI: 10.1515/ling.2000.016.
- Tomei, Renato & Andrea Hollington. 2018. Jamaican linguistic practices in Africa: Music as creative space of language. In *22nd Biennial Conference: “Connecting the Caribbean: Languages, Borders, and Identities” of the Society for Caribbean Linguistics*, 5–11. Costa Rica, Heredia & Limón.
- Vigouroux, Cécile B. 2013. Francophonie. *Annual Review of Anthropology* 42(1). 379–397. DOI: 10.1146/annurev-anthro-092611-145804.
- Warnke, Ingo Hans. 2017. Postcolonial language studies. In Dirk Göttsche, Axel Dunker & Gabriele Dürbeck (eds.), *Handbuch postkolonialismus und Literatur*, 96–100. Stuttgart: Metzler.
- Wells, John Christopher. 1982. *Accents of English. Vol. 3: Beyond the British Isles*. Cambridge: Cambridge University Press.
- Winer, Lise. 1990. Orthographic standardization for Trinidad and Tobago: Linguistic and sociopolitical considerations in an English Creole community. *Language Problems and Language Planning* 14(3). 237–268.

Part II

Partially restructured varieties in the circum-Caribbean region

Chapter 5

A Gbe substrate model for discontinuous negation in Spanish varieties of Chocó, Colombia: Linguistic and historical evidence

Eliot Raynor

Indiana University, Bloomington & Princeton University

In this paper I highlight the central role of an oft-neglected source language model in the presence of discontinuous negation, also known as NEG2 (e.g. *yo no sé cuál es no* ‘I don’t know which one it is’), in the Spanish of Chocó, Colombia. Namely, I point to the presence of analogous negation patterns in Ewe, a widely-spoken variety of Gbe across modern-day Ghana, Togo, and Benin. Data from the Trans-Atlantic Slave Trade Database as well as a 1759 census of the enslaved African(-descendant) population in Chocó mining camps strongly suggest that Gbe languages were spoken by one-third of enslaved Africans taken to the Pacific lowland region of Colombia via the Caribbean port of Cartagena in the 17th and 18th centuries. Traditionally, non-canonical patterns of negation in Ibero-Romance contact varieties involving discontinuous (i.e. pre-verbal and utterance-final) and strictly utterance-final negator morphemes have been attributed to Bantu languages such as Kikongo, some varieties of which feature similar structures. However, for the years in which enslaved Africans were being trafficked directly from West Africa to Chocó via Cartagena (1650–1800), archival sources indicate that those from Bantu-speaking regions of West Central Africa comprised close to one-sixth of the total African-born population, and thus just around half as many as the Gbe-speaking group. Together these findings lead to the conclusion that speakers of Gbe (especially Ewe) played a foundational role in the development of non-canonical negation patterns in Spanish varieties spoken in Chocó, Colombia.



1 Introduction

The specific origins of non-standard linguistic structures in varieties of European languages spoken by majority African-descendant populations in the Americas are the subject of persistent inquiry and much debate among scholars of Atlantic creoles as well as other Caribbean and mainland contact varieties of European languages. The present analysis centers on the presence of one such structure, referred to here as discontinuous negation (abbrev. NEG2), in varieties of Spanish spoken in the Pacific lowland region of northwestern Colombia, exemplified in (1–4).

- (1) (Ruíz-García 2001)¹
Yo **no** sé un nombre **no**
1SG NEG know a name NEG
'I don't know (of) a name.'
- (2) (Ruíz-García 2001)
Yo **no** voy a subir allá **no**
1SG NEG go to go.up there NEG
'I'm not going to go up there.'
- (3) (Ruíz-García 2001)
Esse niño **no** me habla a mí **no**
That child-DIM NEG me talks to me NEG
'That kid doesn't talk to me.'
- (4) (Schwegler 1991a: 109)
Mi papá **sí no** fue a la escuela **no**
My father EMPH NEG went to the school NEG
'My father definitely didn't go to any school.'

The prototypical NEG2 pattern consists of the pre-verbal negative morpheme *no* plus a duplicate utterance-final *no* with no perceivable prosodic break. The seamless transition from the end of the propositional content and the utterance-final negator morpheme can be seen in Figure 1.

1.1 Discontinuous negation in Ibero-Romance

The use of NEG2 in varieties of Chocó Spanish is of sufficient salience that it has often been reported as a distinguishing feature of this variety, for instance

¹All examples from Ruíz-García (2001) were extracted directly from that author's corpus of tape-recorded conversational speech collected in and around Tadó, Chocó.

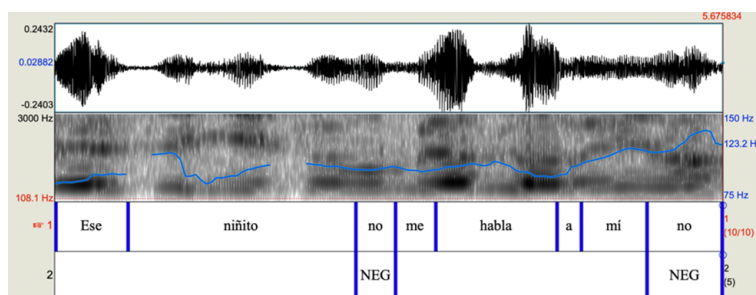


Figure 1: Pitch contour of NEG2 token from example (3)

in the work of Colombian dialectologists (Flórez 1950, Granda 1988, Montes Giraldo 1974). Analogous patterns have been documented as well for other Ibero-Romance contact varieties, including Dominican Spanish (Meggenney 1990, Ortiz López 2007, Schwegler 1991b) and Brazilian Portuguese (Schwegler 1991d, Schwenker 2005). Significant to the present analysis is the fact that discontinuous negation has been discussed extensively in the literature on creole languages, in particular Spanish-based Palenquero (Dieck 2000, Schwegler 1991c, 2016, 2018) and Portuguese-based Santome (Ferraz 1979, Güldemann & Hagemeyer 2019, Hagemeyer 2008, 2009), Angolar (Maurer 1995), and Fa d'Ambô (Post 2013).

The use of NEG2 in negated propositions is not categorical in Chocó Spanish lects, however, since it appears alongside and with less frequency than the canonical Spanish pre-verbal negation structure (NEG1), as in (5–7).²

- (5) (Schwegler 1991a: 96)
 No me gustó y me vine
 NEG me pleased and me came
 ‘I didn’t like it (there), and I came (here).’
- (6) (Schwegler 1991a: 107)
 Yo no me caí
 1SG NEG me fell
 ‘I didn’t fall.’

²Not discussed here is a third option (NEG3) in which only the utterance-final negative morpheme is used, e.g. *ellos llegan aquí no* ‘they don’t come here’ (Ruíz-García 2001: 102). While NEG3 is a robust negation strategy in other Ibero-Romance contact varieties such as Brazilian Portuguese, Palenquero, and Principense, this variant is especially infrequent in Chocó Spanish. Schwegler (1991a) found only one NEG3 example and Ruíz-García (2001) found just three in a speech corpus of 37 hours; thus while its existence is noted, its import is considered marginal.

- (7) (Schwegler 1991a: 111)
No hubo cura para él
NEG was priest for him
'There wasn't a priest for him.'

While a lengthy discussion of the degree of markedness of NEG2 lies outside the scope of the present discussion, informants from my own fieldwork in the capital city of Quibdó described the pattern as stigmatized and yet still *muy común* 'very common' and *inconsciente* 'unconscious'.

1.2 Discourse-pragmatic considerations

The factors governing the variation between NEG1 and NEG2 in Chocó Spanish are at present not well understood and have not been discussed in detail in previous work; this is largely due to the fact that no balanced sociolinguistic corpus of Chocó speech exists at present. Pragmatics-based analyses of variable negation strategies in Brazilian Portuguese (e.g. Schwenter 2005) suggest that NEG2 is licensed only in contexts where the negated proposition is mentioned or "activated" in the immediately preceding discourse. However, in an analysis of 66 tokens of NEG2 extracted from spontaneous conversational data from Chocó, Ruíz-García found that only 54.6% of cases involved negation of a proposition activated in the prior discourse (i.e. old information), while slightly less than half (45.4%) negated new, "out of the blue" propositions (2001: 109–123). A comparison of these cases is demonstrated in examples (8–9), which reveals the immediate discourse preceding the use of NEG2 from examples (2) and (3) above.

- (8) (Ruíz-García 2001)
- Speaker A: ¿Y ustedes ya lo fueron a ver a él allá?
and 2PL already him went to see to he there
'And you already went to see him there?'
- Speaker B: Yo no voy a subir allá no.
1SG NEG go to go.up there NEG
'I'm not going to go up there.'
- Speaker C: ¿Quién llegó abuela?
who arrived grandmother
'Who's here, grandma?'
- Speaker D: Ese niño no me habla a mí no
that child-DIM NEG me talks to me NEG
'That kid doesn't talk to me.'

In (8), B's negated proposition 'I'm not going to go up there' is activated in A's question as to whether B 'already went to see him there'. However, the negation by D relates to the proposition 'X doesn't talk to me', which is not directly activated by C's question about who has just arrived. Thus, the use of NEG2 in Chocó Spanish does not appear to involve the same pragmatic conditioning that has been proposed for Brazilian Portuguese, but rather constitutes an alternative to NEG1 licensed across a wide range of discourse-pragmatic contexts.

1.3 Goal and outline

The objective of the present study is to shed light on the plausible origins of NEG2 in Chocó Spanish, departing from an analysis of the sociodemographic data available regarding the African-descendant population in the Pacific lowlands from the beginnings of significant gold-mining operations in the late 17th c. through the boom and eventual collapse of the slave-labor-based economy in the region in the late 18th century. Section 2, below, presents a sketch of the sociohistorical context in which Chocó Spanish emerged, which can be broadly characterized as one of vast demographic disproportion, wherein the European(-descendant) population never numbered greater than 2.5% of the total population. Section 3 then outlines the patterns of negation in each plausible substrate language – that is, any language that the historical record suggests would have been the L1 of non-negligible numbers of speakers during the relevant time period (1650–1800). Finally, Section 4 weighs the contribution of the most likely substrate candidates in light of the historical evidence and the presence or absence of NEG2-like negation patterns in those languages, concluding that the large proportion of L1 Gbe speakers present in the early stages of Chocó's transformation to a central mining hub of western Colombia provided an analogous and plausible source language model for NEG2. A numerically less prominent yet substantial group of L1 Kikongo speakers likely helped to reinforce the NEG2 pattern, such that it became entrenched in some of the Afro-Hispanic contact varieties that emerged in Chocó.

2 Social and historical considerations

Chocó is one of the thirty-two *departamentos* 'provinces/states' of Colombia, an administrative status it has held since 1944. In the most recent nationwide census, Chocó's total population numbered 388,476, the vast majority of which (82.1%) self-identified as *negro o afrocolombiano* 'black or Afro-Colombian', while the

second largest group (12.7%) self-identified as *amerindio* ‘Amerindian’; significantly, only 5.2% of Chocó’s residents self-identified as *blanco o mestizo* ‘white or mixed white/Amerindian ancestry’ (DANE 2010). These percentages contrast starkly with nationwide demographic figures, in which the *blanco o mestizo* category comprises 86% of the total Colombian population, dwarfing that of both *negro o afrocolombiano* (10.6%) and *amerindio* (3.4%).

2.1 Demographics of 18th c. Chocó

An overview of the history of Chocó shows that similar demographic trends have existed since the earliest large-scale censuses of the region were prepared in the mid- to late-18th century.

Table 1: 18th c. demographics of Chocó by ethnic group

Year	“Whites”	“Slaves”	<i>Libres</i>	“Indians”	Total
1778	332	5,756	3,160	5,414	14,662
1779	335	5,916	3,348	5,693	15,292
1781	336	6,557	3,612	6,202	16,707
1782	359	7,088	3,899	6,552	17,898
1808	400	4,698	15,184	4,450	24,732

Table 1 above, adapted from Sharp (1976: 199), shows that in terms of proportions of the total population, the European-descendants (“Whites”) category reached its maximum in 1778 at just 2.26%, steadily dropping to 1.62% by 1808. On the other hand, the percentages of enslaved Africans and African-descendants (“Slaves”) and Amerindians (“Indians”) are roughly equal throughout the same time period, the former reaching a maximum of 39.3% in 1782 and the latter peaking at 37.2% in 1779 before both decreased both in terms of raw numbers and percentages in 1808, which saw a massive increase in the number of *libres de todos los colores* ‘freed people of all colors’. While the factors leading to this drastic rise in the number of *libres* remain unclear, contemporary Colombian historian Claudia Leal has suggested some possible explanations:

[T]he Afro-descendant population (slaves plus *libres*) increased from 61.4 percent of the total to 80.6 percent, while Indians decreased from 36.6 percent to 17.8 percent. The extraordinary rise in the number of *libres* is hard to explain. It could be justified if slaves and Indians changed categories to

libres with the same rates of their decrease in numbers. The purchase of freedom could explain how some individuals moved from slaves to libres, but no similar explanation exists for Indians becoming libres. (Leal 2018: 50)

The decrease in the Amerindian population very well may reflect an increased tendency towards marronage, a process noted by Werner Cantor (2000) as a common method of escape from semi-captivity employed by the indigenous population of Chocó. This is better understood in light of the fact that the category “Indians” as tabulated in 18th and early 19th century censuses almost certainly referred *only* to tribute-paying Amerindians forced to live in *corregimientos* and provide labor for mine-owners (e.g. riverine transportation, food production, and housing construction), while a proportion of the total Amerindian population maintained their traditional ways of living outside the bounds of Spanish control.

The sociodemographic data from Table 1 point to two groups from which any substrate features of Chocó Spanish must be derived: African(-descendant)s and Amerindians. Tracing the linguistic background of the Amerindian population in Chocó is straightforward, as only Emberá- and Wounaan-speaking communities have inhabited this region continuously³ from the first contacts with Spanish colonists in the early 16th c. through the present day (Jiménez Meneses 2004, Williams 2005). While serious discussion of the Amerindian influences on Chocó varieties of Spanish has unfortunately been omitted in previous accounts, socio-historical factors including the size of the population, coupled with relative linguistic homogeneity within the Chocoan language family (Loewen 1957, 1960, 1963), lend weight to the importance of considering these languages when discussing issues related to language contact in Chocó.

On the other hand, it must be acknowledged that the vast majority of the current population of Chocó is of African descent, as noted above. Prior analyses and the author’s own fieldwork suggest that this population is currently monolingual in what has been described as a “partially restructured” variety of Spanish (Ruíz-García 2001) reflective of “advanced second language acquisition” at an earlier stage in the development of the variety (Sessarego 2016, 2019). Indeed, even as

³The qualifier “continuously” is important here, since Cuna-speaking communities once resided in Chocó as well. Historical accounts by Werner Cantor (2000), Williams (2005), and Vargas Sarmiento (1993) point out that tensions between the Emberá and Cuna populations of Chocó, exacerbated by early contacts with Spanish and English colonists, caused the Cuna to permanently relocate north to Darién, Panama, along the border with Colombia, as well as to the San Blas islands off the Panamanian coast.

early as 1816, two entries from the diary of a Spanish military officer on an expedition to root out revolutionaries in Chocó suggest that the African-descendant population of Chocó spoke target-like Spanish:

May 19: [W]e came across two blacks, and they said the enemies had spent the afternoon prior in the hills above... May 20: [F]rom some blacks that we found along the way we know that the enemies have fled in great haste.⁴

In order to highlight possible substrate sources of NEG2, a variety of sources are considered below to discern the plausible linguistic backgrounds of the largest groups within the African-born and African-descendant population that was trafficked to Chocó between the late 17th c. and 18th century.

2.2 Origins of enslaved Africans arriving in Cartagena, 1650–1800

Cartagena was the primary port of entry of enslaved Africans into colonial New Granada (modern-day Colombia) and thus is the relevant place to look for archival information concerning their places of origin. Regarding Chocó in particular, William F. Sharp states that “most of the miners in the Chocó during the last decades of the seventeenth century were *bozales* [African-born] and most came from Cartagena by way of the Atrato” (Sharp 1976: 111).⁵ Sharp also cites archival accounts from as late as 1777 demonstrating mine owners’ preference for “newly imported blacks who were best suited for work in the mines” (117).

Table 2 summarizes data from the Trans-Atlantic Slave Trade database showing the provenance of 13,371 enslaved Africans sold in Cartagena during the time period spanning 1650–1800 (Voyages Database 2009). The range of locations from which Africans were purchased was vast, stretching from the French-held posts of Cacheu and Saint-Louis in Upper Guinea in the north all the way

⁴Transcription and translation here are mine of an original archival document located in the “Latin American mss.— Colombia, 1558–1890” collection at the Lilly Library in Bloomington, Indiana.

⁵Sessarego argues that “it would be a mistake to assume that blacks in Chocó were all *bozales* who spoke African languages” (2017: 40) and suggests instead that *criollo* [American-born] slaves were more prominent in Chocó. In doing so, he highlights the sales of *criollos* in southeastern city of Popayán, pointing out that most mine (and slave) owners were from there. While these are relevant considerations, it must be noted that a total of only 1,074 *criollos* were sold in Popayán over the course of the entire century (1690–1789) in which mining peaked in Chocó (Colmenares 1997 [1979]: 36). For comparison, if we set aside local birth and death rates, these *criollos* would have constituted just 12% of the entire African(-descendant) population of Chocó in 1778.

5 *A Gbe substrate model for discontinuous negation*

Table 2: Broad regions of origin of Africans disembarking in Cartagena (late 17th-18th c.)

Time Period	Upper Guinea	Sierra Leone	Gold Coast	Bight of Benin	Bight of Biafra	Central Africa	Cape Verde	Príncipe Island
1651–1675	556			966	1,088	650		
1676–1700	820		131	1,138		1,078	313	1,005
1701–1725	245	403	1,801	1,955	344	720		
1776–1800					158			
Total	1,631	403	1,932	4,059	1,590	2,448	313	1,005
%	12.1%	3.0%	14.5%	30.4%	11.9%	18.3%	2.3%	7.5%

south to West Central Africa (present-day Angola, Congo) and including all major regions of the West African slave trade in between. Notably these include the islands of Cape Verde and Príncipe, which served as entrepôts for the shipment of slaves to the Americas, and where Portuguese-based creole languages had already emerged prior to this time period.⁶ On the other hand, the largest percentage (30.4%) arrived from the Bight of Benin, also often referred to as the Slave Coast, a region which Eltis & Richardson (2010: 110–112) illustrate would have stretched from contemporary eastern Ghana to western Nigeria and thus would have covered areas occupied by speakers of Gbe languages, especially Ewe and Fon – for a dialect map of contemporary varieties of Gbe, see Capo (1991: xxiii–xxiv). From this broad region the database shows that more than two-thirds (2,899) of enslaved Africans purchased in the Bight of Benin and arriving in Cartagena were purchased from a single location, Whydah,⁷ between the years 1676–1725. Those arriving in the earlier period between 1651–1675 primarily came from “Ardra”, also known as “Allada”.⁸ This half-century span of time corresponds fairly closely with the earliest period in which it is known that

⁶While slaves arriving from each of these islands comprised relatively smaller groups (especially Cape Verde), their presence is noteworthy, given the claims of McWhorter (1999, 2006) that small numbers of slaves speaking African-origin pidgins or creoles were responsible for catalyzing the creation of creoles across the Atlantic.

⁷“Whydah” corresponds to the modern-day city Ouidah in Benin (for a social history of Ouidah, see Law 2004).

⁸Aboh (2015: Ch.2) provides a helpful overview of the history and significance of the kingdom of Allada, the center of which was a city of the same name situated 50 kilometers inland in modern-day Benin. This is the source of the ethnic denomination Arará, which was applied to enslaved Africans arriving from this location (as seen below).

large numbers of African-born slaves were purchased in Cartagena and transported to Chocó for work in the nascent gold mining industry.

Considerably less specific are the details provided by the Trans-Atlantic Slave Trade database concerning the places of purchase of the second-largest group (18.3%) of Africans disembarking in Cartagena, labelled simply “West Central Africa and St. Helena, port unspecified”. Thus, we can state with some degree of certainty only that this group would have been broadly comprised by speakers of Bantu languages, among them certainly Kikongo, which was central to the formation of Palenquero in the maroon community of San Basilio de Palenque just inland from Cartagena (Schwegler 2002, 2011). Similarly vague is information about Gold Coast arrivals. These comprised the third-largest group (14.5%) of slaves arriving in Cartagena, for which a specific place of purchase, Cape Coast Castle, is known only for 426 of the 1,801 enslaved Africans arriving between 1701–1725, all others falling under the headings “Gold Coast, port unspecified” and “Gold Coast, Fr. definition”. For the purposes of this analysis, then, we will make the general (and rather commonplace) assumption that many among this population would have been Akan/Twi-speaking.

2.3 Origins of enslaved Africans in Chocó, 1759

Some of the uncertainty inherent in the assumptions made above is mitigated by the existence of a census dated 1759 providing information about the enslaved peoples working in fifty-eight mining camps across the region of Chocó at the time.⁹ Particularly illuminating from this document are lists of the names of enslaved Africans and African-descendants working in each camp. What follows below is a summary of the data gleaned from a detailed tabulation of the names found in this original handwritten document.

Out of 2,741 named¹⁰ individuals, 1,535 consist of just a first name, while 1,206 provide a second element consisting of either a surname, an African ethnic denomination, or some type of descriptor. Each of the descriptive second elements

⁹Accessed in the online database of the Archivo General de la Nación de Colombia (Sección Colonia, Negros y Esclavos, Cauca, Legajo 4) <http://www.archivogeneral.gov.co>. It should be noted that Granda (1988) analyzed the same document and reached similar conclusions. However, the numbers reported here do not match precisely those summarized in Granda’s article, and do not invoke a priori assumptions, such as general groupings of diverse ethnonyms under a single linguistic family label (e.g. Kwa).

¹⁰The total population of Chocó in 1759 was 3,918. The difference between the overall total and the 2,741 accounted for here are individuals described collectively as *muleques* ‘children’, *viejos* ‘elderly’, and/or *chusma* ‘unfit for work’.

of the individuals' names was labelled with a code designating subtypes of naming conventions. The range of subtypes included African ethnic denominations (e.g. *Bambará*, *Lucumí*), the generic terms *Criollo* and *Bozal*, surnames of Spanish origin (e.g. *García*, *Figueroa*), racial descriptors (e.g. *Mulato*, *Negro*), other descriptors (e.g. *Cabezón* 'big-headed', *Bailador* 'dancer'), toponyms/demonyms (e.g. *Cartagena*, *Panameño*), marital status (e.g. *Casado* 'married', *Soltero* 'single'), and others which were not immediately discernible (e.g. *Cuco*, *Pino*). Table 3 gives an overview of the subtypes used in the second elements of names found in the census. Most evident is the largest group (43.9%) described in terms describing a range of African ethnicities, the most frequent of which are listed in Table 4.

The predominance of the names *Mina* (137 individuals) and *Congo* (85 individuals) suggest that the linguistic backgrounds associated with these ethnic denominations would have been likely candidates as West African substrate languages in 18th c. Chocó. In the past, historians have tended to associate *Mina* with the Gold Coast port of Elmina in modern-day Ghana. However, Hall points out that

[i]n the Portuguese, Spanish, and French colonies, the 'Mina' were the *casta* Mina, and they are to be identified with the Gbe language speakers as described by Alonso de Sandoval in both the 1627 and 1647 versions of his

Table 3: Subtypes of descriptions following names in 1759 census of Chocó

African ethnic denomination	Criollo ^a	Spanish surname	Race	Demonym	Other ^b
529	405	123	67	10	72
43.9%	33.6%	10.2%	6.0%	0.8%	5.5%

^aIt is noteworthy that such a large proportion of individuals are described as *Criollos*, given the evidence discussed in Section 2.2 that this population would have been considerably smaller than the African-born population in Chocó. An important consideration here is that the percentages reported in Table 3 were calculated in reference to the total of just those individuals whose names are accompanied by a second element in the 1759 census. Thus, if we consider all 2,741 named individuals, the names followed by *Criollo/Criolla* comprise a considerably smaller percentage (14.8%) of the total. Further, since the term *criollo* was typically applied to African-descendant slaves who were born in the colony and spoke Spanish, we may fairly safely assume that few of the remaining 1,535 individuals who were not described explicitly as *Criollo* would have been in fact *criollos*, since this aspect of their identity would have been apparent to the census preparers upon demonstration of their linguistic abilities.

^bSubsumed in the 'Other' category are those elements which do not provide any hints as to the provenance of the person. Besides the elements which were undiscernible as mentioned above (totaling 39 individuals), these also included the term *Bozal* (3 individuals), marital status (8 individuals), and other descriptors (22 individuals).

Table 4: African ethnic denominations in 1759 census of Chocó

Mina	Congo	Arará	Carabalí	Chambá	Chalá	Setre	Mandinga	Popó	Other ^a
137	85	47	46	28	26	25	21	18	96
25.9%	16.1%	8.9%	8.7%	5.3%	4.9%	4.7%	4.0%	3.4%	18.1%

^aOther' subsumes forty-seven distinct ethnic denominations, most appearing only once or twice (e.g. *Fori*, *Bran*), while a few others, such as *Nangó* (13 individuals) and *Tembo* (11 individuals), appear more frequently.

book. He distinguished the *casta* Mina from the Popos, Fulaos, Ardas or Araraes, although he considered that all were related and they were all one (*que todo es uno*). Sandoval makes it clear that the Mina *casta* is to be identified with Gbe-speaking Africans of the Bight of Benin, and specifically the Ewe, Aja, Fon and others of the Gbe language group, and not immigrants from the Gold Coast. (Hall 2003: 70–71)

Law has called into question the definitive nature of Hall's claim about the meaning of *Mina*, pointing out that in many cases

the aggregation of peoples who were linguistically distinct but geographically adjacent (in Africa) are best explicable on the assumption that many people in these groups were bilingual, so that smaller groups could be assimilated into larger ones in the Americas. (Law 2005: 267)

In light of this, it is relevant to note that Mosquera (2008: 84) provides archival testimony from a mining settlement in Tadó, Chocó, in 1728 displaying the multilingual abilities of an individual named Antonio Mina, who it appears was able to serve as an interpreter for other enslaved Africans of various ethnolinguistic backgrounds, as indicated by their names: José Nongo, Marcos Chalá, and Francisco Arará.

In any case, the high frequency of appearance of *Minas* in the 1759 census, coupled with the presence of significant numbers of both *Arará* (47 individuals) and *Popó* (18 individuals), suggests that the Gbe-speaking and/or Gbe-descendant¹¹

¹¹This is an important consideration, since second-generation African-descendants were likely among those who still identified as *Mina*. Indeed, the surname *Mina* (as well as others, e.g. *Carabalí*) is still found in African-descendant communities from the Pacific lowlands of Colombia and Ecuador.

population recorded in the document reaches a total of 202 individuals, or 38.2% of those described in terms of African ethnic denominations.

In the case of *Congo*, which appears in 18.1% of the names from the 1759 census, there is little room for debate as to the provenance of these individuals, particularly given the close correspondence with the data in Table 2 showing that 18.3% of the slaves arriving in Cartagena came from West Central Africa. I will again assume (as in Section 2.2) that members of this group were speakers of Bantu languages, especially Kikongo. As for other ethnic denominations appearing in the 1759 census, in the interest of clarity and a degree of certainty, here no specific claims are made as to their linguistic backgrounds.¹²

3 Negation strategies in candidate substrate languages

In this section, I present the morphosyntactic structures used to express negation in each of seven likely candidate substrate languages highlighted above in Section 2: Emberá, Wounaan, Ewe, Gen, Fon, Kikongo, Akan/Twi, and Principense. These negation strategies are illustrated in examples (9–20) below.

3.1 Amerindian languages

I begin with two indigenous languages of the Chocoan language family, given that these (especially Emberá) were the most widely spoken for the longest time period in Chocó, i.e. from initial contacts with Spanish-speaking colonists in the early 16th c. through to the present. As illustrated in examples (9–12), three varieties of Emberá as well as Wounaan express negation in the form of suffix attached to the verb stem, which can optionally be followed by additional suffixes expressing tense, mood, or aspect.

- (9) Emberá Pedee (Harms 1994: 131)

mi warra k^bäi-ʔe

my son sleep-NEG

‘My son is not sleeping.’

- (10) Emberá Katío (Mortenson 1999: 98)

ünũ-ẽ-pa-s^hi-a

see-NEG-EQ-PAST-DECL

‘He didn’t see her.’

¹²For instance, *Carabalí* appears frequently but can be traced to a range of languages in modern-day Nigeria.

- (11) Emberá Chamí (Aguirre Licht 1999: 112)
ídjí kúi-wē-a
today bathe-NEG-DECL
'Today I haven't taken/won't take a bath.'
- (12) Wounaan (Holmer 1963: 129)
mə yeka-ba-m
1SG talk-NEG-FUT
'I'm not going to talk.'

Examples (9–12) demonstrate that each of the Chocoan languages spoken by indigenous peoples in contact with Spanish in Chocó has a similar typological profile with respect to the expression negation, that is, morphologically by way of verbal suffixation. As such, neither Emberá nor Wounaan make particularly good candidates for the introduction of NEG2 into Chocó Spanish, since no pre-verbal element is present in any variety.

3.2 West African languages

Next I turn to Ewe and Fon, the two varieties of Gbe most likely to have been present in Chocó during the early stages of transformation of this region into a major center of the gold mining industry in the 18th century. Examples (13–14), drawn from 19th c. and 20th c. grammars of Ewe, demonstrate that negation in this variety involves the placement of a free morpheme *me* before the verb – as well as before any TMA markers, as seen in (13) – plus the morpheme *o* at the end of the sentence, following any post-verbal complements or adjuncts. Crucially, both pre-verbal *me* and sentence-final *o* are obligatory in the expression of negation in these varieties.

- (13) Ewe (Ellis 1890: 234)
Nye **me** a du **o**
1SG NEG FUT eat NEG
'I will not eat.'
- (14) Ewe (Ameka 1991: 64)
Kofi **me** va afi sia **o**
Kofi NEG come place this NEG
'Kofi did not come here.'

Similarly, Gen negation requires the expression of both pre-verbal *mú* and utterance-final *ò* negator morphemes (Aboh 2004: 46–47).

On the other hand, the Fon variety of Gbe can express negation pre-verbally with the morpheme *ma* or post-verbally with the sentence-final morpheme *a*, but according to Aboh (2004: 45), the pre-verbal negator *ma* and post-verbal *a* cannot co-occur – unless they are in a conditional clause introduced by *ni* ‘if’, in which case they must both be present. See (15–16) for examples.

- (15) Fon (Aboh 2004: 44)
 Kɔku **ma** na xɔ asɔn le
 Koku NEG FUT buy crab NUM
 ‘Koku will not buy the crabs.’
- (16) Fon (Aboh 2004: 45)
 Kɔku na xo asɔn le **a**
 Koku FUT buy crab NUM NEG
 ‘Koku will not buy the crab.’

Generally, then, while Ewe, Gen, and Fon have some similarities in terms of negation through the use of free morphemes appearing pre-verbally and/or sentence-finally, only Ewe and Gen appear to have a pattern of negation that could have provided a substrate source for NEG2 in Chocó Spanish.

Continuing within the Niger-Congo language family, examples (17–18) below are from Kikongo. Given the historical data presented above and prior proposals for the origin of NEG2 in Chocó (e.g. Schwegler 1991a, 2018), this language is considered particularly relevant as a possible substrate. On the other hand, Bantu languages in general, and Kikongo in particular, present the most complex range of possibilities of any of the languages considered here in terms of negation strategies. *The World Atlas of Linguistic Structures* describes Kongo as the only language¹³ “with an obligatory clause-final negative word, a choice between an immediately preverbal negative word or a negative prefix, and an optional suffix” (Dryer 2013).

Examples (17–18) present cases in which Kikongo negation partially aligns with that of the NEG2 pattern found in Chocó Spanish, but in both cases the pre-verbal negator is attached as a prefix to the verb and followed by a series of morphemes for person/number and tense/aspect also prefixed to the verb. Furthermore, in (17) the post-verbal negator appears before an adverb and prepositional phrase, whereas in Chocó Spanish the post-verbal negative morpheme in NEG2 follows those same types of constituents (see examples 2 and 4, above).

¹³Here and below names for this language are presented in the sources that have been cited. In all cases it is clear that varieties of the same language (Kikongo) are being described.

Thus, while NEG2-like constructions are possible, the number of variable strategies of negation in Kikongo and the complexities of its morphosyntactic realization do not make it an ideal candidate as a substrate for the NEG2 structure as used in the Spanish of Chocó.

- (17) Kisikongo
ki-a-mon-idi o npangi aku ko mazono ku zandu
NEG.1SG-PAST-see-PRF AUG brother your NEG yesterday at market
‘I didn’t see your brother yesterday at the market.’
- (18) Kongo (Lumwamu 1973: 200)
ka-t-a-dí-di-éti ko
NEG-1PL-COMPL-eat-PAST-REL.COMPL NEG
‘We have not yet eaten.’

While not as likely a candidate in light of the sociohistorical profile of Chocó, Akan/Twi should be considered as well, since languages of the Kwa branch of Niger-Congo are often lumped together and considered typologically similar in language contact work (see, e.g., Granda 1988). Akan/Twi negation, as in (19) below, involves the use of a verbal prefix – in this case preceded by an aspectual marker – with no post-verbal marking whatsoever.

- (19) Akan/Twi (Amfo 2010: 104)
Papa no a-n-ko adwuma nnera
man DEF COMPL-NEG-go work yesterday
‘The man didn’t go to work yesterday.’

Similar to the Chocoan languages, then, Akan/Twi is an unlikely source model for NEG2 in the Spanish of Chocó.

Finally, to conclude the discussion of possible substrate models for NEG2, example (20) below comes from the Portuguese-based creole Principense. While there is no definitive evidence to suggest that speakers of any Portuguese-based creole were trafficked to Chocó, it is possible that speakers of Principense may have been among the 1,005 enslaved Africans that arrived in Cartagena via Príncipe between 1676–1700, some of whom may have indeed been taken to the Pacific lowland region.

- (20) Principense (Maurer 2009: 58)
Te ninge nhon di pasa lala fa
have person no of pass there NEG
‘There is nobody who passes by over there.’

However, as the sentence in (20) shows, Principense uses only a sentence-final negator *fa*¹⁴. The morpheme *nhon* glossed as ‘no’ above has scope only over the word *ninge* ‘person’, combining to mean ‘nobody’ as seen in the English translation. Thus, Principense is also not a strong candidate as a source for the appearance of NEG2 in Chocó.

3.3 Summary

Table 5 summarizes the types of negation strategies found in the seven candidate substrate languages considered most relevant to the context in which Chocó Spanish emerged. The NEG2 pattern presented by Chocó Spanish involves the use of pre-verbal *and* sentence-/utterance-final negator morphemes; therefore, the most likely candidate in purely linguistic terms is Ewe, which presents an analogous pattern. Kikongo may also be of relevance given the *possibility* of analogous structure, with the caveats discussed above.

Table 5: Types of negation among candidate substrate languages

Language	Subgroup (Family)	Negation type
Emberá	Chocoan (Amerindian)	Verbal suffix
Wounaan	Chocoan (Amerindian)	Verbal suffix
Ewe	Gbe (Niger-Congo)	Pre-verbal and sentence-final morpheme
Gen	Gbe (Niger-Congo)	Pre-verbal and sentence-final morpheme
Fon	Gbe (Niger-Congo)	Pre-verbal or sentence-final morpheme
Kikongo	Bantu (Niger-Congo)	Pre-verbal morpheme or verbal prefix and/or suffix <i>and clause-final morpheme</i>
Akan/Twi	Tano (Niger-Congo)	Verbal prefix
Principense	Portuguese-based creole	Sentence-final morpheme

4 Conclusion

The present analysis is narrow in scope both in terms of the linguistic structure under analysis (discontinuous negation or NEG2) as well as the geographical re-

¹⁴This is in fact unique to Principense as compared to the closely-related Portuguese-based creoles Santome, Angolar and Fa d’Ambô. According to Maurer (2013), “[t]he fact that ... Principense has no double negation like Santome *na ... fa* is probably due to the existence of the validator *na* (epistemic modality) in Principense, which has the same shape and the same position of the negator *na*, namely immediately preceding the tense, aspect, and mood markers”.

gion of its use (Chocó, Colombia). However, the account offered here is perhaps somewhat more far-reaching insofar as NEG2 has been the subject of some debate regarding its origins in Chocó as either a contact feature derived from a West African source (Granda 1978: 514, 1988, Ruíz-García 2001, Schwegler 1991a) or one that developed across Ibero-Romance in the 15th c. via Jespersen's cycle and thus reflecting Spanish varieties spoken by colonists (Sessarego 2017). The latter argument has been challenged and largely discounted by Schwegler (2018) in part because the pattern of NEG2 in Chocó as seen in examples (1–3) above involves a pre-verbal and a sentence-final negator after objects, adverbs, and other intervening morphemes, while Jespersen's cycle¹⁵ involves immediately post-verbal negators, as do all of the examples in Sessarego's (2017) analysis of 15th–19th c. Spanish from the Davies (2002) *Corpus del Español*. There is no existing evidence of NEG2 occurring in colonial varieties of Spanish, nor has any attempt been made to provide data showing its use in western Colombia during the relevant time period. On the other hand, there is an abundance of evidence that NEG2 appears frequently in Ibero-Romance varieties in intense contact with West African languages, including Afro-Hispanic varieties of Chocó, Colombia and the Dominican Republic, as well as Brazilian Portuguese and the creole languages Palenquero, Santome, Angolar, and Fa d'Ambô. The origin and persistence of NEG2 in each of these varieties requires an analysis based on local sociohistorical and linguistic evidence.

In the case of Chocó it has been particularly important to adopt a local perspective, as opposed to a macroscopic or comparative one. The structure of the present paper thus proceeds from a review of original archival documents that shed light on the languages most likely to have been in the right place at the right time to have exerted an influence on Spanish in Chocó – in accordance with Bickerton's Edict – before considering them as possible substrates. The historical and linguistic evidence I present here indicates that speakers of Gbe varieties, in particular Ewe and Gen speakers enslaved in the Bight of Benin and sold in Cartagena in the late 17th c. and early 18th c., were in all likelihood the key agents that introduced NEG2 structures found in contemporary varieties of Spanish spoken in Chocó. My account allows room for the contribution of speakers of Bantu languages as well, particularly Kikongo, who were enslaved in West Central Africa and trafficked through Cartagena in the relevant time period, but whose numerical presence was substantially less than that of the Gbe-speaking group and

¹⁵This is the version of the cycle as originally conceived by Jespersen (1917). Recently, van der Auwera (2009) has offered an expanded version of the cycle that includes the development of patterns with sentence- or utterance-final negator morphemes – notably, this has been invoked primarily to account for these sorts of patterns in West African languages, not European ones.

whose language does not present negation patterns strictly analogous to NEG2. Speakers of Chocoan languages (especially Emberá) were the only other group present in Chocó in the 17th and 18th c. with large enough populations to have significantly influenced the development of Chocó Spanish. It is apparent that NEG2, however, did not originate in any of these varieties, all of which express negation by way of verbal suffixation.

Acknowledgements

In addition to two anonymous reviewers and the editors of this volume, several people deserve specific recognition for generously offering their time and insight, from which this analysis has benefited enormously. First and foremost, I have to thank two of my mentors at Indiana University, J. Clancy Clements and Kevin J. Rottet, for their engagement with and encouragement of this line of inquiry since its humble beginnings over three years ago at the time of writing. Secondly, I owe a great deal of credit for the present analysis to Martha Ruíz-García, whose dissertation planted the seeds for this project, and whose invaluable corpus I was given the opportunity to digitize in 2018. I am also indebted to a number of participants who attended my talks at the 2017 and 2019 summer meetings of the Society for Pidgin and Creole Linguistics and Associação de Crioulos de Base Lexical Portuguesa e Espanhola, where I had the opportunity to discuss this and other matters related to the Spanish varieties of Chocó, Colombia. Among many who have been kind enough to correspond and explore/debate these questions with me, I'd like to thank in particular Miguel Gutiérrez Maté, Danae Perez, Luana Lamberti, Johan van der Auwera, Armin Schwegler, and Sandro Sessarego. Of course, any inaccuracies herein remain my own.

References

- Aboh, Enoch Oladé. 2004. *The morphosyntax of complement-head sequences: Clause structure and word order patterns in Kwa*. Oxford: Oxford University Press.
- Aboh, Enoch Oladé. 2015. *The emergence of hybrid grammars: Language contact and change*. Oxford: Oxford University Press.
- Aguirre Licht, Daniel. 1999. *Embera*. Munich: LINCOM Europa.
- Ameka, Felix. 1991. *Ewe: Its grammatical constructions and illocutionary devices*. Canberra, Australia: Australian National University. (Dissertation).

- Amfo, Nana Aba Appiah. 2010. Negation in Akan: Linguistic convention versus pragmatic inference. *Journal of Asian and African Studies* 80. 97–112.
- Capo, Hounkpati B. C. 1991. *A comparative phonology of Gbe*. Berlin: Foris Publications.
- Colmenares, Germán. 1997. *Historia económica y social de Colombia II*. Bogotá: Tercer Mundo Editores.
- DANE. 2010. *Departamento administrativo nacional de estadística - análisis regional de los principales indicadores sociodemográficos de la comunidad afrocolombiana e indígena a partir del censo general 2005*. Bogotá: Revista Escala.
- Davies, Mark. 2002. *Corpus del español: 100 million words, 1200s-1900s*. %7Bhttp://www.corpusdelespanol.org/hist-gen/%7D.
- Dieck, Marianne. 2000. *La negación en palenquero: Análisis sincrónico, estudio comparativo y consecuencias teóricas*. Madrid: Iberoamericana.
- Dryer, Matthew S. 2013. Position of negative morpheme with respect to subject, object, and verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of linguistic structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/144>.
- Ellis, Alfred B. 1890. *The Ewe speaking peoples of the Slave Coast of West Africa*. London: Chapman & Hall.
- Eltis, David & David Richardson. 2010. *Atlas of the transatlantic slave trade*. New Haven: Yale University Press.
- Ferraz, Luis I. 1979. *The creole of São Tomé*. Johannesburg: Witwatersrand University Press.
- Flórez, Luis. 1950. El habla del chocó. *Thesaurus* 6. 110–116.
- Granda, Germán de. 1978. *Estudios lingüísticos hispánicos, afrohispanicos y criollos*. Madrid: Gredos. 522.
- Granda, Germán de. 1988. Los esclavos del chocó: Su procedencia africana (siglo XVIII) y su posible incidencia lingüística en el español del área. *Thesaurus* 43. 65–80.
- Güldemann, Tom & Tjerk Hagemeyer. 2019. The history of sentence negation in the Gulf of Guinea Creoles. *Journal of Ibero-Romance Creoles* 9. 55–84.
- Hagemeyer, Tjerk. 2008. Double-headed negation in Santome. *Journal of Portuguese Linguistics* 7. 62–83.
- Hagemeyer, Tjerk. 2009. Aspects of discontinuous negation in Santome. In Norbert Cyffer, Erwin Ebermann & Georg Ziegelmeyer (eds.), *Negation patterns in West African languages and beyond*, 139–166. Philadelphia: John Benjamins.
- Hall, Gwendolyn M. 2003. African ethnicities and the meaning of ‘Mina’. In Paul E. Lovejoy & David Vincent Trotman (eds.), *Trans-Atlantic dimensions of ethnicity in the African diaspora*, 65–81. New York: Continuum.

- Harms, Phillip Lee. 1994. *Epena Pedee syntax*. Dallas: Summer Institute of Linguistics.
- Holmer, Nils M. 1963. Gramática comparada de un dialecto del chocó (con textos, índice y vocabulario). In Statens Humanistika Forskningsråd (ed.), *Estudios chocoes*, 79–248. Gothenborg: Etnografiska Museet.
- Jespersen, Otto. 1917. *Negation in English and other languages*. Copenhagen: Høst.
- Jiménez Meneses, Orián. 2004. *El chocó: Un paraíso del demonio. Nóvita, citarà y el baudó, siglo XVIII*. Medellín: Editorial Universidad de Antioquia.
- Law, Robin. 2004. *Ouidah: The social history of a West African slaving ‘port’, 1727–1892*. Athens, OH: Ohio University Press.
- Law, Robin. 2005. Ethnicities of enslaved Africans in the diaspora: On the meanings of “Mina” (again). *History in Africa* 32. 247–267. DOI: 10.1353/hia.2005.0014.
- Leal, Claudia. 2018. *Landscapes of freedom: Building a postemancipation society in the rainforests of western Colombia*. Tuscon: University of Arizona Press.
- Loewen, Jacob A. 1957. *Comparative dictionary: Choco dialects-Spanish-English*. Cali, Colombia: Mennonite Brethren Board of Missions. 600.
- Loewen, Jacob A. 1960. Dialectología de la familia lingüística Chocó. *Revista Colombiana de Antropología* 9. 9–22.
- Loewen, Jacob A. 1963. Chocó I: Introduction and bibliography. *International Journal of American Linguistics* 29. 239–263.
- Lumwamu, François. 1973. *Essai de morphosyntaxe systématique des parlers kongo*. Paris: Klincksieck.
- Maurer, Philippe. 1995. *L’angolar. Un créole afro-portugais parlé à São Tomé. Notes de grammaire, textes, vocabulaire*. Hamburg: Buske.
- Maurer, Philippe. 2009. *Principense: Grammar, texts, and vocabulary of the Afro-Portuguese Creole of the Island of Príncipe, Gulf of Guinea*. London: Battlebridge.
- Maurer, Philippe. 2013. Principense. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *The survey of pidgin and creole languages*, vol. 2: Portuguese-based, Spanish-based, and French-based languages, 72–80. Oxford: Oxford University Press.
- McWhorter, John H. 1999. The afrogenesis hypothesis of plantation Creole origin. In Magnus Huber & Mikael Parkvall (eds.), *Spreading the word: The issue of diffusion among the Atlantic Creoles*, 111–152. London: University of Westminster Press.
- McWhorter, John H. 2006. Creole transplantation: A source of solutions to resistant anomalies. In Clancy Clements, Thomas A Klingler, Deborah Piston-Hatlen & Kevin J Rottet (eds.), *History, society and variation: In honor of Albert Valdman*, 103–133. Amsterdam: John Benjamins Publishing Company.

- Megenney, William W. 1990. *África en Santo Domingo: Su herencia lingüística*. Santo Domingo: Editorial Tiempo.
- Montes Giraldo, Joaquin. 1974. El habla del chocó: Notas breves. *Thesaurus* 39. 409–428.
- Mortenson, Charles A. 1999. *A reference grammar of the northern Embera languages*. Dallas: SIL International. 208.
- Mosquera, Sergio A. 2008. *Mondongo: Etnolingüística en la historia afrochocoana*. Bogotá: Arte Láser.
- Ortiz López, Luis A. 2007. La negación en la frontera dominico-haitiana: Variantes y usos (socio)lingüísticos. In Kim Potowski & Richard Cameron (eds.), *Spanish in contact: Policy, social and linguistic inquiries*, 211–233. Amsterdam/Philadelphia: John Benjamins.
- Post, Marike. 2013. Fa d'Ambô structure dataset. In Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath & Magnus Huber (eds.), *Atlas of Pidgin and Creole structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://apics-online.info/contributions/38>.
- Ruíz-García, Martha. 2001. *El español popular del chocó: Evidencia de una reestructuración parcial*. Albuquerque, NM: University of New Mexico. (Dissertation).
- Schwegler, Armin. 1991a. El habla cotidiana del chocó (Colombia). *América Negra* 2. 85–119.
- Schwegler, Armin. 1991b. La doble negación dominicana y la génesis del español caribeño. *Lingüística* 3. 31–88.
- Schwegler, Armin. 1991c. Negation in Palenquero: Synchrony. *Journal of Pidgin and Creole Languages* 6. 165–214.
- Schwegler, Armin. 1991d. Predicate negation in contemporary Brazilian Portuguese: A linguistic change in progress. *Orbis* 34. 187–214.
- Schwegler, Armin. 2002. On the (African) origins of Palenquero subject pronouns. *Diachronica* 19. 273–332.
- Schwegler, Armin. 2011. Palenque(ro): The search for its African substrate. In Claire Lefebvre (ed.), *Creoles, their substrates, and language typology*, 225–249. Amsterdam/Philadelphia: John Benjamins.
- Schwegler, Armin. 2016. Truth reset: Pragmatics in Palenquero negation. In Armin Schwegler, John McWhorter & Liane Ströbel (eds.), *The Iberian challenge: Creole languages beyond the plantation setting*, 231–267. Madrid: Iberoamericana.
- Schwegler, Armin. 2018. On the controversial origins of non-canonical Spanish and Portuguese negation: Case closed? *Lingua* 202. 24–43. DOI: 10.1016/j.lingua.2017.12.004.

- Schwenter, Scott A. 2005. The pragmatics of negation in Brazilian Portuguese. *Lingua* 115(10). 1427–1456. DOI: 10.1016/j.lingua.2004.06.006.
- Sessarego, Sandro. 2016. On the non-(de)creolization of Chocó Spanish: A linguistic and sociohistorical account. *Lingua* 184. 122–133. DOI: 10.1016/j.lingua.2016.06.005.
- Sessarego, Sandro. 2017. Chocó Spanish double negation and the genesis of the afro-Hispanic dialects of the Americas. *Diachronica* 34(2). 219–252. DOI: 10.1075/dia.34.2.03ses.
- Sessarego, Sandro. 2019. *Language contact and the making of an Afro-Hispanic vernacular: Variation and change in the Colombian Chocó*. Cambridge: Cambridge University Press.
- Sharp, William F. 1976. *Slavery on the Spanish frontier: The Colombian chocó, 1680-1810*. Norman, OK: University of Oklahoma Press.
- van der Auwera, Johan. 2009. Chapter 3. The Jespersen cycles. In Elly van Gelderen (ed.), *Linguistik Aktuell/Linguistics today*, vol. 146, 35–71. Amsterdam: John Benjamins Publishing Company. DOI: 10.1075/la.146.05auw.
- Vargas Sarmiento, Patricia. 1993. *Los embera y los cuna: Impacto y reacción ante la ocupación española en los siglos XVI y XVII*. Bogotá: Instituto Colombiano de Cultura.
- Voyages Database. 2009. *Voyages: The Trans-Atlantic Slave Trade Database*. <https://www.slavevoyages.org> (30 October, 2019).
- Werner Cantor, Erik. 2000. *Ni aniquilados, ni vencidos: Los emberá y la gente negra del atrato bajo el dominio español*. Bogotá: Instituto Colombiano de Antropología e Historia.
- Williams, Caroline A. 2005. *Between resistance and adaptation: Indigenous peoples and the colonization of the Chocó 1510-1753*. Liverpool: Liverpool University Press.

Chapter 6

The minutes of the *Irmandade Nossa Senhora do Rosário dos Homens Pretos* in Recife: A case study

Fernanda M. Ziober

Vrije Universiteit Amsterdam

This article is a case study of the written texts by Manoel de Barros, a member of the *Irmandade Nossa Senhora do Rosário dos Homens Pretos* ‘Brotherhood of Our Lady of the Rosary of Black Men’ in the 19th century, in the city of Recife, Brazil. The minutes and terms of the brotherhood constitute important sociolinguistic documentation of literacy of people with limited access to texts and formal education. The texts do not follow the writing standards of 19th century Portuguese. The variations from the standard norm have three main sources: (i) orthographic, (ii) phonetic and (iii) morphosyntactic. Characteristics such as apheresis and epenthesis are compared to those occurring in similar texts written by the African and Afro-Brazilian people of Bahia (Oliveira 2006), on Bahian Afro-Brazilian Portuguese (Lucchesi et al. 2009) as well as Portuguese in general.

1 Introduction

This article is a case study of the linguistic characteristics of the historical writings of Judge-Brother Manoel de Barros of the *Irmandade de Nossa Senhora do Rosário dos Homens Pretos* in the 19th century.¹ The *irmandades*, ‘brotherhoods’, were civil organizations that allowed their members to secure a space for Catholic religious worship, a burial place, and other amenities. In view of the fact that these books have been contaminated by bacteria, I want to give visibility to such

¹“Brotherhood of Our Lady of the Rosary of Black Men”, a Catholic brotherhood of African men and African descendants.



documentation as a source of historical sociolinguistic data that deserves documentation and digitization. In Brazil, after the publication of Law 10.639/03 which stipulates the mandatory teaching of the history and culture of Afro-Brazilians throughout the entire educational system, there has been a growing interest in this historical topic and the subsidizing of new pedagogical materials. Related to this issue, there has also been interest in the formation and stabilization of the national variety of Portuguese referred to as Brazilian Portuguese.

According to Mattos e Silva (2004) and Noll (2008), the slave trade and the dispersion of enslaved individuals across the Brazilian territory were fundamental factors for the implementation and diffusion of the Portuguese language in Brazil. Up until then, competence in Standard Portuguese was restricted to Portuguese descendants and people who had the financial means to study in the metropolis. Given the process of acquisition of a second language during the colonial period,² the aforementioned authors argue that the African and Afro-Brazilian speakers of Portuguese were agents in the changes and innovations that occurred in Brazilian Portuguese compared to European Portuguese.³ However, there are few linguistic studies based on historical data of Afro-Brazilian and African writers, like manuscripts and large-scale written corpora, to confirm the hypothesis proposed by Mattos e Silva (2004).

Generally speaking, it is very difficult to locate such manuscripts, whether personal or institutional writings. This is because, in addition to the unfavourable historical socioeconomic conditions in Brazil, the enslaved population was prevented from learning to read and write and, more specifically, they were institutionally forbidden to attend public classes by the Constitution of 1824, although no specific penalty has been reported for offenders. Indeed, there are reports and records of enslaved people present in schools as well as people who knew how to read and write at various levels of proficiency (Silva 2007).

In my search for manuscripts, I came across several studies concerning the Catholic brotherhoods *dos homens pretos* 'of black men', which were important institutions for the African and Afro-Brazilian population during the slavery period. There are two research areas that study these institutions: History (e.g., Assis 1998, Levi 2006, Mac Cord 2005; Quintão 2002) and Linguistics (cf. Oliveira 2006, Galves & Lobo 2019). A number of these studies focus on the states of Rio de Janeiro, Minas Gerais and Bahia. Fortunately, as Pernambuco was one of the

²In Brazilian linguistic literature, Lucchesi et al. (2009) define this phenomenon as an irregular linguistic transmission of the Portuguese language.

³There are characteristics of Brazilian Portuguese that are considered as representing a conservative variety of European Portuguese (cf. Graebin 2017, and cf. Massini-Cagliari et al. 2016, on comparative synchronic studies in Portuguese segmental phonology).

states with the highest number of abolitionists and had many freed former slaves already in the 17th and 18th centuries (Klein 1969), a large number of 19th century manuscripts have survived in its brotherhoods. This documentation has proven to be a relevant source for understanding the history of Africans and their descendants, not only in Brazil, but also the rest of the world.⁴

This is a case study identifying characteristics of popular urban Portuguese, as spoken in Recife, more specifically in the neighbourhood of Santo Antonio. Although this brotherhood functioned in the geographical and institutional centre of the African and Afro-Brazilian community, I do not consider the variety of Portuguese featured in the texts an ethnic variety, but part of a popular and urban regional variety in a multicultural and multi-ethnic context. As the volume of data is not sufficient to allow for a statistical analysis, I present the documentation, the historical context and the fundamentals for using it as a database in this article.

After a brief introduction, the article proceeds with a short history of the brotherhoods, their operational structure and historical importance for the African and Afro-Brazilian population during that period, including their role as an institution that granted access to literacy. Next, I discuss the methodology and the palaeographic criteria essential for the selection of data. After evaluating three books from the brotherhood and the transcription of one thereof, I compared the signatures and the handwriting of the terms and minutes contained therein. In assessing the sociolinguistic nature of these materials, the ethnic background of Manoel de Barros was confirmed because of his position as a judge within the brotherhood, a position that could only be held by significant individuals in the Afro-Brazilian community. The section poses some challenges, such as the classification of the textual genre of the minutes.

In the third section I compare my data to (1) Oliveira (2006) mainly for phonological phenomena; (2) Lucchesi et al. (2009) for a comparison with the characteristics of Afro-Brazilian Portuguese, such as the optional plural inflection on nouns; and (3) Petter (2009) for Portuguese in general. Finally, I present pertinent summary comments and suggest possibilities for future research involving the documentation and transcription of such materials.

2 Materials and methods

From the outset of the colonial period, Catholic religious brotherhoods were present in every location where Africans were disembarked. The first records

⁴I also found information about these minutes and terms in Cape Verde and Portugal.

point to the emergence of the Brotherhood of Our Lady of the Rosary in Europe,⁵ the first one being in the German city of Düsseldorf, in 1407. For the black brotherhoods, on the other hand, there are records of the *Confraria da Nossa Senhora do Rosário dos Homens Pretos* 'Confraternity of Our Lady of the Rosary of Black Men' in Cape Verde, still in the 15th century and the *Irmandade de Nossa Senhora do Rosário dos Homens Pretos* 'Brotherhood of Our Lady of the Rosary of Black Men' in Angola in the 16th century.⁶ The creation of a brotherhood was marked by the creation of a statute. As soon as the brothers began to congregate, they needed their statute to be officially approved by the metropolis (Lisbon). Once approved, they could raise funds to build a church with a cemetery, and for this reason many foundation dates of churches might be considered as the foundation date of a brotherhood.

The brotherhoods were civil organizations that derived from the liberal organizations of artisans of the European Middle Ages. In the colonies, they permitted the enslaved and freed population to secure a space for Catholic religious worship, a burial place (since the cemeteries at that time were private and located in churches attended by the white community), and they provided other amenities, such as cash loans and health care. Assis (1998) argues that the brotherhoods *dos homens pretos e pardos* 'of black and brown men' occupied an ambiguous space in society, because on the one hand they acted as an agent of assimilation and control, since members could not have participated in revolts or have been arrested, yet, on the other hand, they constituted the only institution of assistance to this population. Quintão (2002) observes that the greatest advantage of such associations was their autonomy, a fact that contributed to Brazilian religious syncretism.⁷

The statute that governed the brotherhoods and confraternities defined the rules of operation and membership. Usually, an initial membership fee was charged, and to occupy positions of responsibility, such as those of judge and clerk,

⁵In this case, we are not talking about the brotherhoods *dos homens pretos* 'of black men', but the European brotherhoods *da Nossa Senhora do Rosário* 'of the Rosary' which gave origin to the brotherhoods *da Nossa Senhora do Rosário dos Homens Pretos e Pardos* 'of the Rosary of black and brown men'.

⁶There are records from the 17th century onwards concerning the emergence of both the aforementioned institutions in Brazil. The first brotherhoods *da Nossa Senhora do Rosário dos Homens Pretos e Pardos* 'of the Rosary of black and brown men' were the one in Rio de Janeiro in 1639, the one in Belém do Pará in 1682, the ones in Salvador and Recife in 1685 and the one in Olinda and Igarassu in 1688.

⁷Considering that this topic goes beyond the scope of this article, I suggest reading Juruá (2011), and the literature on *candomblé*, a syncretic Afro-Brazilian religion, and *Umbanda*, a syncretic religious practice combining Afro-Brazilian and Brazilian Amerindian ritual features.

it was necessary to have social prestige. In the case of the brotherhood of the Santo Antonio do Recife neighbourhood, according to the appointment regulation (*compromisso*) in force at the time, it was also necessary for the King of Congo,⁸ elected by the brotherhood itself, to be a descendant of the indigenous people of Angola (*Compromisso 1782*),⁹ which binds this brotherhood to its Angolan origin.¹⁰ However, one cannot say that all members were from the Angola region or that this would have been the case for their descendants since people from different parts of Africa were brought to Recife.

An issue that has caused much historical controversy is the fact that, due to limited literacy in general and, possibly, due to the institutional prohibition of access to education for the African and Afro-Brazilian population, the majority of the Catholic brotherhoods *dos homens pretos e pardos* ‘of black and brown men’ in Pernambuco hired white clerks to write their minutes. The justification for Manoel de Barros, the judge of the brotherhood who had already been clerk of the same brotherhood, for writing the deliberations (*termos*), was the temporary financial difficulty of the brotherhood, which prevented it from hiring a clerk, as was the tradition and, in a way, the obligation of such institutions. Moreover, in the documents under consideration, the signatures that were preceded by the expression “Cruz de...” [‘Cross of...’] were actually written by the clerk and the cross functioned as the signature, while those who knew how to write their own names would sign in their own hand, as shown in the two figures below.

Silva (2007) does an excellent job of reconstructing the presence of people from the African and Afro-Brazilian community in Pernambucan society. Using data from the 1872 census,¹¹ Silva (2007) compares the number of literate people in the enslaved population in Pernambuco, which was 1 literate person for every 999 illiterate ones and 0.6 literate enslaved children for every 999.4 illiterate ones. However, it is possible that some enslaved people did not openly admit their literacy, since that could have been considered dangerous by the authorities.

I analysed the advertisements of “Escravos fugidos”¹², “Avisos diversos”¹³ and

⁸In Recife, the King of Congo was an important public figure who interceded alongside official bodies in the representation of the freed and enslaved population.

⁹[G]entio de Angola” (*Compromisso 1782*).

¹⁰In 1491, the *manicongo* ‘King of the Congo and Angola’ Nzinga-a-Nkuwu, also called Nkuwu Nzinga, was the first African king to convert to Catholicism, being baptized as John I, having his kingdom recognized by the Pope and even sending one of his sons to be trained in the Vatican (Souza 2002). Despite this, the same king deconverted himself later.

¹¹Cf. Brasil (1876).

¹²‘Runaway slaves’.

¹³‘Miscellaneous notices’.

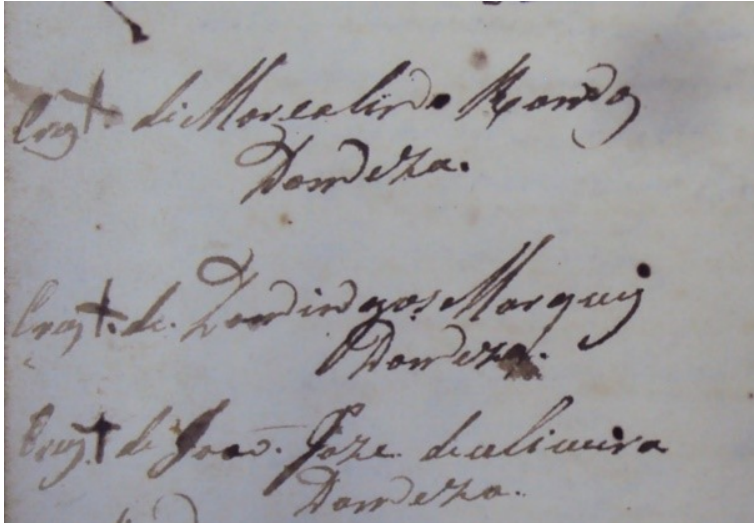


Figure 1: Signatures made by crosses



Figure 2: Handwritten signature

“Vendas”¹⁴ from the newspaper *Diário de Pernambuco*¹⁵ between the years 1831 and 1848¹⁶. In 342 advertisements, I found two references to two literate enslaved men, Agostinho and Joaquim. When discussing alternative forms of literacy, Moysés (1994) reports that some enslaved people and freedmen used shopping lists, booklets and even a copy of their freedom letters as a way of learning to read and write.

In this way, Manoel de Barros, the clerk who after three years became a judge, was probably a freedman. This allowed him to properly exercise his duties as

¹⁴‘Sales’.

¹⁵‘Pernambuco Daily’.

¹⁶Most of them were from the 1840s and all the advertisements are stored in the Jordão Emerenciano State Public Archive.

a person of prestige in the black community and someone who had access to literacy.

With regard to methodology, I adopt, in general, a qualitative analysis. I visited several archives in Pernambuco and found the relevant documentation in possession of IPHAN.¹⁷ The *Irmandade Nossa Senhora do Rosário dos Homens Pretos* had donated its records to that institution for the purpose of secure storage and conservation. By means of digital photography, copies were made of three books, out of which I transcribed only one, containing information dated between 1829 and 1833. For the transcription, I used the paleographic techniques of Barbosa et al. (2006).

Altogether, 32 texts were fully transcribed, 13 of which were written by Manoel de Barros. The remaining texts were in three different calligraphies. Although these very likely were African or Afro-Brazilian clerks, I chose to do a case study of the linguistic characteristics of Manoel de Barros as he held the position of judge and clerk and signed his own name. Consequently, I was sure that the clerk was writing in his own hand (Figure 3).¹⁸

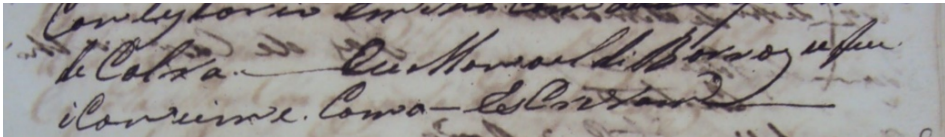


Figure 3: Handwritten signature

After selecting the texts for analysis, I conducted a paleographic transcription of the documents and identified the main fixed textual structures of the minutes, such as title, place and date. The stylistic rigidity of the texts minimized the syntactic possibilities therein in view of the fact that several phrases were always repeated, for example *Estando reunidos em mesa...* ‘Being together at the table...’ or *Aos três dias do mês...* ‘On the third day of the month...’. Perhaps, because of such repetitions, copying the minutes themselves may have been a means of improving the writing skills of the author.

One major challenge was the lack of comparisons between texts written by African and Afro-Brazilian clerks on the one hand and texts written by other

¹⁷Instituto do Patrimônio Histórico e Artístico Nacional (lit. National Institute of Historical and Artistic Heritage).

¹⁸The excerpt below reads: *Eu Manoel de Barros o fiz e assinei como Escrivão* ‘I Manoel de Barros recorded and signed as clerk’.

Brazilian clerks, for example *brancos* ‘white men’ and *pardos* ‘brown men’.¹⁹ Until the beginning of the 19th century, the literate population was minimal, and most clerks were either Portuguese or had lived in Portugal. Even so, I compared our data with previous linguistic analyses of minutes written by Africans and Afro-Brazilians in Salvador, Brazil, and there were several similarities. I also took into consideration the reported characteristics of Afro-Brazilian Portuguese, as described by Lucchesi et al. (2009),²⁰ and Portuguese varieties in general.

Despite the hypothesis presented by Mattos e Silva (2004), there are characteristics from this last-mentioned dialect that are shared by speakers of Popular Brazilian Portuguese (PBP) in general, so the question remains whether the characteristics described are exclusive to the Afro-Brazilian community or whether these characteristics also extend to PBP regardless of the ethnic origin of the speakers. Therefore, considering that the Portuguese written by Manoel de Barros is not an isolated variety, I treat the Portuguese written by him not as Afro-Brazilian Portuguese, but as part of PBP from Recife, which was written by an Afro-descendant. Hence, I understand the relationship between the two varieties that, although belonging to the same ethnic group, developed in a different way.

3 Linguistic characteristics of the terms/minutes

The description of the linguistic characteristics of the texts written by the African/Afro-Brazilian clerk Manoel de Barros is not intended to be a generalization of the work produced by all Afro-Brazilian clerks of the time, but to compare the characteristics of this clerk’s texts to the writings of other Afro-Brazilian and African clerks (Oliveira 2006), to the linguistic characteristics attributed to the Afro-Brazilian Portuguese (Lucchesi et al. 2009) and to Brazilian Portuguese (Petter 2009). The main goal of this discussion is to demonstrate the potential of the documents of the brotherhoods and confraternities of *Nossa Senhora do Rosário dos Homens Pretos* in order to reach an understanding of the linguistic interaction between speakers in urban environments in the 19th century.

From a total of 32 texts, I analysed three types of characteristics: (1) phenomena of an orthographic nature; (2) phonetic phenomena, which I distinguished between (a) those that occur at the level of the phonological word, and (b) those

¹⁹Although I performed the paleographic transcription of texts by other brothers from the same period, in this work I was unable to compare their characteristics. Moreover, in this research, I only analysed the linguistic characteristics already attributed to the Afro-Brazilian and African population in Brazil.

²⁰The authors described the so-called “Afro-Brazilian Portuguese”, the variety of Portuguese spoken by four rural Afro-Brazilian communities in the state of Bahia.

that occur at the level of more than one phonological word, namely, hypo-and hyper-segmentation; and (3) morphosyntactic phenomena. I redefine some of the phenomena listed in Ziober (2014) and below I exemplify these points. Some phenomena are recurrent, such as the elevation of the vowel <e> to <i>, with 79 occurrences, while other occurrences happen only once, such as the rhotacism from <l> to <r>, in <sipruco> *sepulcro* ‘sepulchre’.

3.1 Phenomena of a graphic nature

First, I need to acknowledge graphic phenomena as phenomena resulting from variation in spelling. Compare the variation between <s>, <c>, <ç>, <ss> and <sc> for the phoneme /s/:

- (1) variation between S, C, Ç, SS and SC
 - a. *acima* → <açima> PREP ‘above’
 - b. *segundo* → <çegundo> NUM ‘second’
 - c. *posse* → <poçia> NOM ‘the act to take office’
 - d. *fosse* → <fusce> VERB ‘were/was’
 - e. *essa* → <esa> DEM ‘this’
 - f. *obrigação* → <obriga com> NOM ‘obligation’
 - g. *senhor* → <cinhor> PRON ‘Sir’
 - h. *do nascimento* → <donascimento> PREP-DET NOM ‘of the birth’
 - i. *o nosso* → <onoco> DET PRON ‘our’

I found (2) two orthographic possibilities, <ch> and <x>, for the palatal /ʃ/:

- (2) variation of CH and X
 - a. *se acha* → <ciaxa> PRON VERB ‘find it’
 - b. *Paixão* → <Paichao> PNOM ‘Paixão’ (lit.Passion)

Because of articulatory reasons, the grapheme <m> is consistent before the bilabials <p> and , and <n> before other consonants. Nevertheless, at that time, such norms were not yet widespread and Manoel de Barros presented variation in the use of these segments, even within the same text.

- (3) variation of M and N
 - a. *dezembro* → <dezenbro> NOM ‘December’
 - b. *consistório* → <Comcistorio> NOM ‘religious counsel’

Next, Manoel de Barros favoured the use of <u> instead of <v> in consonantal contexts, but never the contrary. Two hypotheses present themselves: first, that cursive writing has similar shapes, making it difficult to recognize them in handwriting; second, that the clerk deliberately wrote forms very close to <u> because of his partial knowledge of the orthographic norm. Thus, writing both almost the same could be a way of not incurring in spelling errors.

- (4) variation of U and V
festividade → <festiuidade> NOM ‘festivity’

Finally, I found other graphic phenomena occurring only once, namely, variation in the spelling of <h>, <j>, <rr>, <qu> and <f>.

3.2 Phonetic phenomena

I believe several phonological phenomena are phonetically conditioned. At first, I present data at the level of the phonological word. The use of <z> in contexts between vowels is related to both graphic and phonetic reasons. Because the spelling is not intuitive to the writer, words are written with <s> although the pronunciation of the sibilant sounds like [z]:

- (5) use of <z> in contexts of [z]
a. *mesa* → <meza> NOM ‘table’
b. *Rosário* → <Ruzario> PNOM ‘Rosary’

The first phenomenon of purely phonological motivation is metathesis. In their study, da Hora et al. (2007) define the metathesis of Portuguese in the 18th century as a transposition of consonants in one or two syllables. According to the authors, metathesis involving liquid consonants is of the tautosyllabic perceptual type, because the syllables alternate between one syllable where the liquid is in coda position, or perceptually two syllables where the liquid occurs in a branched onset. This phenomenon has been observed in historical studies of Portuguese (Coutinho 1976, Faria 1970, Oliveira 2006), but not all the data found by Oliveira (2006: 349) seemed to be phonetically motivated. Also, Petter (2009: 162) mentions that the metathesis is common in the Portuguese spoken in Angola²¹ and Mozambique,²² e.g., *pruguntar* < *perguntar* ‘to ask’. So, metathesis, as well as other phenomena, is not exclusive to the documents or region considered, but it has occurred in many other varieties of Portuguese.

²¹Cf. Chavagne (2005).

²²Cf. Laban (1999).

(6) metathesis

- a. *partido* → <pratido> ADJ ‘broken, split’
- b. *aprovada* → <aporuada> ADJ ‘approved’
- c. *congregados* → <comergados> ADJ ‘member of a congregation’

Another phenomenon observed is rhotacism, the transition between lateral and vibrant liquids due to their articulatory similarities. In the data of Oliveira (2006: 418) more than half of the rhotacisms from /r/ to /l/ happened in a coda position. The same occurs in data from Manoel de Barros. In our data, there was only one occurrence of /l/ to /r/, simultaneously with a metathesis, more precisely in the word <sipruco> *sepulcro* ‘sepulchre’.

(7) rhotacism

- a. *contar* → <comtol> VERB ‘to say/to count’
- b. *sepulcro* → <sipruco> NOM ‘sepulchre’

Oliveira (2006: 326) also found aphaeresis, the loss of a word-initial phoneme or syllable. In our data, aphaeresis is found as a result of the graphic omission of some syllables and sounds. In fact, I found an example, still very recurrent in all varieties of Portuguese, (8), i.e. the reduction of the syllable <es> in <estar> *estar* ‘to be’. In addition, I found the total assimilation of the sequence of two vowels [e] between the clitic and the verbal root.²³

(8) aphaeresis

estando → <tondo> VERB ‘staying’

(9) total assimilation

se eleger → <celeger> PRON VERB ‘elect oneself’

A further characteristic observed by Oliveira (2006: 336) is apocope. He notes that it has been a regular phenomenon throughout the history of Portuguese. Petter (2009: 163) also cites it from research by Chavagne (2005) on Angolan Portuguese. The verbal example (10) below coincides with a current characteristic of spoken Portuguese in general, deleting /r/ from the infinitive form of the verb. Furthermore, nowadays, in Recife, verb forms ending with /r/ in the final coda may not be realized or are replaced by an extension of the final vowel.

(10) apocope

querer → <quere> VERB ‘to want’

²³I chose to classify (9) as total assimilation, but since I have only this data, it is hard to judge if it is a case of phoneme loss or assimilation.

Epenthesis, the addition of segments in the middle of a word, also occurs in our data, as in example (11b), possibly motivated by the assimilation of the /l/ in word-final coda position to the stressed syllable and assimilation of the /s/ from the first to the second syllable. Example (11c) features the syllabification of an occlusive in coda with the filling of the nuclear vowel, similar to what happens in Brazilian Portuguese nowadays, for example [advogadʊ] > [adʒivo-gadʊ] ‘lawyer’. Petter (2009: 161) argues that Brazilian, Angolan and Mozambican Portuguese have a preference for an open syllable pattern because of their contact with Bantu languages, in which this pattern is standard. Additionally, various Brazilian indigenous languages have the same standard pattern, which also might have contributed to the epenthesis process. In examples (11a) and (11b), assimilation may also have played a role in the epenthetic process through the presence of the phonemes /d/ (11a) and /ʎ/ (11b).

(11) epenthesis

- a. *reduzida* → <rududiz/ida> ADJ ‘reduced’
- b. *respeitável* → <res/-pestoluel> ADJ ‘respectable’
- c. *insígnia* → <inciguina> NOM ‘badge’

With regard to vowels, the elevation of pre-tonic and post-tonic mid vowels is common in Brazilian Portuguese, and in the case of post-tonic vowels, the elevation is categorical at the end of the word in many Portuguese varieties. Oliveira (2006: 367) observes cases of elevation of stressed mid vowels (15) but does not find similar cases nor an obvious motivation in the literature despite admitting that phonetic motivation is possible and probable. Cases of hypercorrection due to oscillation between <e> and <i> in Portuguese spelling and phonetics are also found in our data. I have two hypotheses for this: (i) I believe examples (13b) and (13c) are due to variation among earlier varieties of Portuguese (Graebin 2017) or (ii) they follow developmental paths similar to Mozambican and Angolan Portuguese, in which [e]>[i] occurs in similar contexts to Brazilian Portuguese, e.g. *minino* ‘boy’, *piquinino* ‘small’ (Mozambican Portuguese) and *piqueno* ‘small’ (Angolan Portuguese).²⁴

(12) elevation of pretonic mid vowels

- a. *de* → <di> PREP ‘of’
- b. *servir* → <ciruir> VERB ‘to serve’
- c. *costume* → <custume> NOM ‘habit’

²⁴For Angolan and Mozambican Portuguese examples, cf., respectively, Chavagne (2005) and Laban (1999).

- (13) elevation of post-tonic mid vowels
- a. *aonde* → <aondi> ADV ‘where’
 - b. *quinze* → <quinzis> NUM ‘fifteen’
 - c. *houve* → <houi> VERB ‘there was’
 - d. *junho* → <junhu> NOM ‘June’

- (14) elevation of stressed mid vowels
- Pessoa* → <Pecua> PNOM ‘Pessoa’ (lit. Person)

A further variation observed in examples (15) is vowel posteriorization. Oliveira (2006: 351) suggests that assimilation processes contributed to such occurrences, and I add that nasalization possibly has interfered with the perception of the vowel, given that some data from Oliveira, as well as all of our data, constituted contexts of nasality.

- (15) vowel posteriorization
- a. *Manoel* → <Monoel> PNOM ‘Manoel’
 - b. *Santíssimo* → <Somticimo> PNOM ‘Santíssimo’ (lit. Sacred)

Oliveira (2006: 371) attributes the lowering of high vowels to the proximity of <e> in the same word; compare (16b) and (16c) for assimilation processes. At the same time, in these examples, I encounter the presence of nasality adjacent to the lowered vowel, showing a possible extension of these assimilation processes beyond vowel quality and nasality. Furthermore, in the case of (16c), assimilation occurs by raising the stressed nasal vowel, followed by its lowering in the process of assimilation.

- (16) lowering of high vowels
- a. *juiz* → <jues> NOM ‘judge’
 - b. *circunstâncias* → <circontoncias> NOM ‘circumstances’
 - c. *de desistência* → <didezentencia> PREP NOM ‘abdication’

Afterwards, I observed diphthong processes, like the ones signalled by Noll (2008) and Oliveira (2006: 409) in Brazilian Portuguese and Petter (2009: 162) in Angolan Portuguese. With regard to the contexts of monophthongization, Oliveira (2006: 408) found a large number of reductions from [ow] to [o], but few cases from [oj] to [o].

- (17) diphthongization
a. *propôs* → <porpois> VERB ‘propose’
b. *toda* → <tuoda> QUANT ‘all’
- (18) monophthongization
a. *dezoito* → <dezoto> NUM ‘eighteen’
b. *outubro* → <otubro> NOM ‘October’

To conclude this discussion of phonetic and phonological variation types, I draw attention to some phonetic phenomena that interact with the morphology and contribute to an understanding of the domain of the morphological and phonological word. Examples (19) and (20) demonstrate the variation in the segmentation of excerpts involving prepositions and clitic pronouns.

In hyper-segmentation, (20a) is a case in which the <a> is segmented from <atual> *atual* ‘current’ because it resembles a definite article *a* ‘the’ before a feminine word, and (20c) segments <nosa> in a similar way from *nossa* ‘our’, although there are two words in the excerpt <nós assinamos> ‘we signed’. Despite the fact that the cases involve word segmentation, I interpret the phenomena in (19) and (20) mainly as graphic ones, demonstrating the phonological character of some morphemes that may contain parts that resemble different words.

- (19) hypo-segmentation
- a. de toda a mesa → <dituadameza>
PREP QUANT DET table
‘from all the members of the table’
- b. e te represento → <eterreperzento>
CONJ PRON represent.1SG
‘and I represent you’
- c. com a data → <comadata>
PREP DET date
‘on the date’
- d. de comum → <dicomum>
PREP common
‘in common’
- e. se eleger → <celeger>
PRON elect
‘elect himself’

(20) hyper-segmentation

- a. *atual* → <a tual> ADV ‘current’
- b. *compromisso* → <com primico> NOM ‘statute’
- c. *nós assinamos* → <nosa cinomos>
1PL sign.1PL
‘we signed’

3.3 Morphosyntactic phenomena

A very widespread phenomenon in Brazilian Portuguese is the variation in plural agreement within the noun phrase, the preference being marking the first element or determiner (Lucchesi et al. 2009). Inverno (2005) also gives many examples for Angolan Portuguese in which the noun is not marked for plural, but the demonstratives, determinants and possessives are, e.g., *Estas duas mulher* ‘these two women’ (lit. ‘these two woman’) and *Os meus filho* ‘my sons’ (lit. ‘the my son’).²⁵ We also observed a similar case in (21a). However, example (21b) shows the confusion these markers caused for the clerk, apparently inducing hypercorrection in the word <quinzis> *quinze* ‘fifteen’, yet at the same time he does not mark the plural on the word *dia* ‘day’.

As the volume of documentation is small and there are many fixed textual structures, a more in-depth study would be necessary to yield concrete clues about the behaviour of the agreement rule, since in several cases plural agreement is marked on all items in the noun phrase.

(21) plural agreement in the noun phrase

- a. *dos* *homen-s preto-s* → <dos homem preto>
PREP-DET-PL man-PL black-PL
‘of Black Men’ (lit. of black man)
- b. *ao-s* *quinze dia-s do* → <aos quinzis diado>
PREP-DET-PL fifteen day-PL PREP-DET
‘On fifteen of...’ (lit. on the fifteen days of...)

4 Discussion and conclusion

This study has documented the main linguistic characteristics of the texts produced by the clerk and judge Manoel de Barros of the *Irmandade Nossa Senhora do*

²⁵These examples are from Inverno (2005) and Petter (2009).

Rosário dos Homens Pretos. Since 2014, such documents have been contaminated by fungi and bacteria and have been removed from public access. Therefore, I would like to appeal for the urgency of restoration, documentation and digitization of this data source, not only in Recife, but in the case of brotherhoods throughout Brazil that experience the same conservation problems.

The brotherhoods are an important source of historical data about African and Afro-Brazilian communities. These brotherhoods allowed people to organize themselves in terms of a civil society and to improve their lives by providing services, loans and amenities at a time when only the upper classes would otherwise have had access to them. Despite the fact that those organizations faced discrimination on a racial basis, there was still a lot of prestige in holding a role as one of the leaders or as a member.

With regard to the linguistic phenomena analysed, four of them were clearly motivated by graphic reasons and sixteen were motivated by phonological aspects of spoken Portuguese. From these, five were related to vowel alteration (12–16), two ((7) and (11)) involved more than one phonological process at the same time and two ((5) and (15)) were related to orthography in addition to the phonological aspect.

Although I encountered phenomena that are frequently found in other Portuguese varieties, e.g. (12), the elevation of pretonic mid vowels, and (17), diphthongization, others are not mentioned in the literature. The elevation of a stressed mid vowel as in (15) and the motivation of this process are less clear.

As far as the texts and their organization are concerned, the fixed textual structures should be considered formal language. For that reason, morphosyntactic phenomena found in them may not constitute as reliable data of spoken Portuguese as those occurring in informal texts, e.g. personal letters and diaries. In that sense, the absence of an official orthography in the late 18th and 19th centuries strengthens the observation of phonologically motivated phenomena and the argument that they are indeed characteristics of spoken Portuguese.

Another important aspect to be mentioned is the fact that these texts were produced in a multicultural environment with many different ethnic backgrounds. Although the institution's origin can be linked to a specific country, in this case Angola, it is not possible to say that all the members were people of Angolan heritage. The most distinctive social aspect at the time was skin colour, so the only distinction we have is between the black/brown brotherhoods and the white brotherhoods, the latter not featuring any explicit mention of skin colour. That said, we cannot claim that the Portuguese spoken by the black/brown brotherhoods shared necessarily more characteristics with Angolan Portuguese than with PBP in general. In order to investigate which linguistic characteristics are

typical of Brazilian Portuguese in general and those which are typical of a certain region or ethnic group, I suggest that this study should be expanded to include not only black/brown brotherhoods.

In conclusion, these texts were produced by urban people and the absence of an official orthography resulted in more transparency of phonological phenomena than the morphosyntactic ones, since those were limited by text typology. Finally, even considering their limitations, these texts are an important data source to understand the society and the writing patterns of less literate people in the 19th century: they are available all over Brazil and Africa and they are one of the few data sources to allow for comparisons between these varieties from a historical perspective.

Acknowledgements

I would like to thank my M.A. thesis supervisors for this research, Stella Telles and Marlos de Barros Pessoa, as well for the members of the brotherhood and for the workers of the archives from Recife, Brazil. I also would like to thank the three anonymous reviewers who helped to significantly improve the text.

Abbreviations

1SG	1 st person singular	DEM	Demonstrative	PNOM	Proper Noun
1PL	1 st person plural	DET	Determinant	PREP	Preposition
ADJ	Adjective	NOM	Nominal	PRON	Pronoun
ADV	Adverb	NUM	Numeral	QUANT	Quantifier
CONJ	Conjunction	PL	Plural	VERB	Verb

Sources

Advertisement of the newspaper *Diário de Pernambuco*. 1831–1846. Transcription of documents from the State Public Archive *Jordão Emerenciano*.

Comprimisso [Statute] *da Irmandade de Nossa Senhora do Rosário dos Homens Pretos da Vila do Recife*. 1782. Arquivo Histórico Ultramarino (AHU–PE), LAPEH, Códice 1303.

Manuscripts of the *Irmandade Nossa Senhora do Rosário dos Homens Pretos do Recife*. 1829–1832.

References

- Assis, Virgínia Maria Almôedode. 1998. *Pretos e brancos – a serviço de uma ideologia de dominação*. Recife, BR: Universidade Federal de Pernambuco. (Thesis).
- Barbosa, Maria do Socorro Ferraz, Vera Lúcia Costa Acioli & Virgínia Maria Almôedo de Assis. 2006. *Fontes repatriadas: Anotações de história colonial*. Recife: Editora Universitária da UFPE.
- Brasil. 1876. *Directoria Geral de Estatística. Recenseamento da população do Imperio do Brazil a quese procedeu no dia 1o de agosto de 1872*. Rio de Janeiro: Typ. De g. Leuzinger e Filhos. <https://archive.org/details/recenseamento1872bras/page/n11/mode/2up>.
- Chavagne, Jean-Pierre. 2005. *La langue portugaise d'Angola – étude des écrits par Université Lumière Lyon 2 rapport à La norma européenne du portugais*. Lyon: Université Lumière Lyon 2. (Dissertation).
- Coutinho, Ismael de Lima. 1976. *Gramática histórica*. 7th edn. Rio de Janeiro: Acadêmica.
- da Hora, Dermeval, Stella Telles & Valéria Neto de Oliveira Monaretto. 2007. Português brasileiro: Uma língua de metátese? *Letras de Hoje* 42(3). 178–196.
- Faria, Ernesto. 1970. *Fonética histórica do latim*. 2nd edn. Rio de Janeiro: Acadêmica.
- Galves, Charlotte & Tânia Lobo. 2019. *O português escrito por afro-brasileiros no século XIX: As atas da sociedade protetora dos desvalidos*. Salvador: EDUFBA.
- Graebin, Geruzade Souza. 2017. A pronúncia das vogais médias no Português europeu e no Português brasileiro: Um estudo diacrônico. *Revista Letras* 96. 104–131.
- Inverno, Liliana C. C. 2005. *Angola's transition to vernacular Portuguese*. Coimbra: Universidade de Coimbra. (MA Thesis).
- Juruá, Padrinho. 2011. *O ritual do Rosário das santas almas beneditas: A presença da Irmandade dos Semiombas e dos Sakáangás na Umbanda*. São Caetano do Sul: Fundação Biblioteca Nacional.
- Klein, Herbert S. 1969. The colored freedmen in Brazilian Slave society. *Journal of Social History* 3(1). 30–52. DOI: 10.1353/jsh/3.1.30.
- Laban, Michel. 1999. *Mozambique: Particularités lexicales et morphosyntaxiques de l'expression littéraire en portugais. Document accompagnant une demande d'habilitation à diriger des recherches*. Paris: Université de la Sorbonne Nouvelle – Paris III.
- Levi, Joseph Abraham. 2006. *Compromisso e solução: Escravidão e as irmandades afro-brasileiras - Origem e Formação das confrarias religiosas no Brasil Colonial (1552-1822)*. Berlin: Lit Verlag.

- Lucchesi, Dante, Alan Baxter & Ilza Ribeiro. 2009. *O português afro-brasileiro*. Salvador: EDUFBA.
- Mac Cord, Marcelo. 2005. *O rosário de d. Antônio: Irmandades negras, alianças e conflitos na história social do Recife, 1848-1872*. Recife: Ed. Universitária da UFPE.
- Massini-Cagliari, Gladis, Luiz Carlos Cagliari & Wayne J. Redenbarger. 2016. A comparative study of the sounds of European and Brazilian Portuguese: Phonemes and allophones. In W. Leo Wetzels, João Costa & Sergio Menuzzi (eds.), *The handbook of Portuguese linguistics*, 56–68. Malden, MA: Wiley Blackwell.
- Mattos e Silva, Rosa Virgínia. 2004. *Ensaio para uma sócio-história do português brasileiro*. São Paulo: Parábola Editorial.
- Moysés, Sarita M. Affonso. 1994. Leitura e apropriação de textos por escravos e libertos no Brasil do século XIX. *Revista Educação e Sociedade* 48(15). 200–212.
- Noll, Volker. 2008. *O português brasileiro: Formação e contrastes*. São Paulo: Editora Globo.
- Oliveira, Klebson. 2006. *Negros e escrita no Brasil do século XIX: Sócio-história, edição filológica de documentos e estudo lingüístico*. Salvador: Universidade Federal da Bahia. (Dissertation).
- Petter, Margarida. 2009. O continuum afro-brasileiro do português. In Charlotte Galves, Helder Garmes & Fernando Rosa Ribeiro (eds.), *África-Brasil: Caminhos da língua portuguesa*, 159–174. Campinas: Editora da Unicamp.
- Quintão, Antonia Aparecida. 2002. *Lá vem meu parente: As irmandades de pretos e pardos no Rio de Janeiro e em Pernambuco (século XVIII)*. São Paulo: Anablume/Fapesp.
- Silva, Adriana Maria Paulo da. 2007. *Processos de construção das práticas de escolarização em Pernambuco, em fins do século XVIII e primeira metade do século XIX*. Editora Universitária da UFPE: Recife.
- Souza, Marina de Mello e. 2002. *Reis negros no Brasil escravista: História da festa de coroação de Rei Congo*. Belo Horizonte: Universidade Federal de Minas Gerais.
- Ziober, Fernanda Maciel. 2014. *O escrivão Manoel de Barros e os termos da Irmandade de Nossa Senhora do Rosário dos Homens Pretos: O português popular do Recife na primeira metade do século XIX*. Recife: Universidade Federal de Pernambuco. (MA Thesis).

Chapter 7

Language contact in Puerto Rico: Documenting an emerging variety of English

Sally J. Delgado

University of Puerto Rico at Cayey

Spanish and English have a unique contact history in Puerto Rico, where English operates in diglossia with vernacular Spanish for most Puerto Rican bilinguals who live on the island. Yet, few speakers recognize a local variety of English, and scholarship on the speech of Puerto Ricans living on the island is limited. This paper tests the hypothesis that speaker involvement in generating a database of linguistic resources on Puerto Rican English can increase awareness of how the Spanish-dominated context of the island has given rise to a distinct variety of English. A language documentation project aimed to test this hypothesis using eight experimental strategies: six data-collection methodologies that documented spoken and written forms of English, and two surveys that measured language attitudes. This paper presents exploratory analysis on three of these data-collection activities relating to: 1) lexical variation in speech, 2) semantic and structural variation in writing, and 3) attitudes on accent and confidence. Findings indicate that Puerto Rican English still tends to be understood and expressed through a Spanish matrix framework, but speakers increasingly assert ownership over their unique variety of English and use it to express Puerto Rican culture and identity. Speaker participation and the involvement of early-career scholars across all eight data-collection activities indicate that these stakeholders are motivated to document and describe their local variety of English. Furthermore, their participation not only resulted in increased recognition of Puerto Rican English but also positive re-evaluation of how local variation in English is manifested through language mixing in the contact situation of the island. The central recommendation of this paper encourages speakers, educators, community leaders and scholars of Puerto Rican studies to document the local variety of English and advocate for its local and international recognition.



1 Introduction

Spanglish is a poorly defined but well-recognized language contact phenomenon among Latino communities (Ardila 2005, Otheguy & Stern 2010, Montes-Alcalá 2018). The name itself, derived from splicing the names of the two languages from which it emerged, embodies a problematic history of colonial conflict and linguistic domination in the Americas that prompted situations of both cultural accommodation and re-appropriation among indigenous populations and second language speakers. The term Spanglish – as the sequence of morphemes that compose its name suggest – has been historically recognized as “[a] type of Spanish contaminated by English words and forms of expression, spoken in Latin America.” (Oxford English Dictionary 1989: v. XVI: 105). Many authors have accordingly recognized the mixture of languages as a problematic and potentially dangerous phenomenon. For example, Salvador Tio describes the state of Spanish in Latin America as follows: “se pudre en la frontera nuevo-mexicana donde [hablan]... la burundanga lingüística” (“[Spanish] is rotting on the New Mexican border where [they speak]... linguistic nonsense”) (Tio, 1992, as cited in Lipski 2004: 1). Likewise, Roberto González-Echeverría describes “spanglish, la lengua compuesta de español e inglés que salió de la calle, [...] plantea un grave peligro a la cultura hispanica y al progreso de los hispanos” (“Spanglish, the language composed of Spanish and English, [...] represents a grave danger for Latino culture and the progress of Latinos.”) (González-Echeverría, 1997, as cited in Lipski 2004: 1–2). In his paper on the linguistic typology of Spanglish, Lipski addressed the controversy and confusion of definitions and recognizes that “[i]n most cases the word Spanglish and the related connotations of linguistic hybridity *qua* illegitimate birth are used to denigrate the linguistic abilities of Hispanic speakers born or raised in the United States” (Lipski 2004: 1). Others have considered the mixed language as a temporary means of communication used by those who do not yet have multilingual competence (Betances Palacios, 1997, as cited in Lipski 2004: 1), or a communicative necessity comparable to a pidgin (Ardila 2005: 78). Although many authors agree that Spanglish is essentially a dialect of Spanish, more neutrally defined as “Spanish characterized by numerous borrowings from English” (American Heritage Dictionary of the English Language 2020), others recognize it as a separate form of speech defined as a “Fusión de español y el inglés” (“fusion of Spanish and English”) (Real Academia Española 2020) created by groups of Hispanics, specifically in the U.S., who mix lexical and grammatical elements of both languages.

Attitudes towards Spanglish as “an ongoing controversy” have been traditionally polemic and often stigmatize the mixed language practice in a U.S. context

(Dumitrescu 2010: 136). However, the use of English and Spanish mixed language expression in the spoken and visual arts, particularly among young people, have been receiving increasing recognition – and, in some respects, appreciation – for the mixed language practices of Hispanic bilinguals in recent decades (Hernández 1997, Stavans 2000, Pousada 2017).¹ Concurrently, as the literature of education explores the value of translanguaging strategies for increasing numbers of migrant children and mixed language communities, interest in understanding how young people mix Spanish and English has gained momentum (Sayer 2008). Interestingly, and regardless of whether the mixture of languages is perceived in linguistic terms as a dangerous threat to Spanish purity, an emerging Spanish dialect, or an intermediary form among bilingual populations, the phenomenon has predominantly been considered as a Hispanic issue. In contrast, the ideological framework of this paper encourages readers to consider the phenomenon of language mixing² from a specifically Puerto Rican and an Anglophone perspective. In full recognition of the ongoing controversy regarding the definition of the mixed language practices commonly known as Spanglish and the perceptions of its many culturally and ethnically distinct bilingual speakers, the scope of this paper does not seek to re-define Spanglish but instead considers how the discourse surrounding it may have affected the recognition of a Puerto Rican dialect of English. As such, this paper engages with the Puerto Rican language contact situation and specifically explores the effects of mixed language practices on how English is spoken and perceived on the island.

This paper does not propose to negate nor diminish the existence of Puerto Rican Spanglish as a separate form of speech. Indeed, international scholarship in contexts of sustained language contact has shown that mixed languages can thrive in multilingual spaces. This is specifically among young speakers who creatively navigate the fluid spaces in between languages, for example, among the predominantly children, adolescents and young adult speakers of Sheng, a mixed language of Nairobi City (Githiora 2002: 176) and among the students who have innovated a new form of Estonian-Russian (Zabrodska 2013). For

¹Defining types of bilingualism is beyond the scope of this study, but I acknowledge the continuum of language competencies that are encompassed by the word “bilingual” (see Pousada 2017: 4–7) and use the term in an inclusive manner for the purposes of this paper.

²This paper defines language mixing as “the merging of characteristics of two or more languages in any verbal communication” (Odlin 1989: 6). In this sense, language mixing practices incorporate a wide range of language contact phenomena including transfer, borrowing, and code-switching in both language acquisition and fluent bilingual contexts (Muysken 2000). Each of these phenomena have their own complexities and sub-classifications and are often discussed as part of an overlapping system of contact features (Romaine & Kachru 1992, Yumoto 1995, Muysken 2000).

many such contact languages popularized among young people, students and their academic institutions often pioneer efforts in documentation and advocacy for recognized status of the new language. For example, the Jamaican Language Unit of the University of the West Indies in Mona, Jamaica, continues to serve a central role in advocacy for Jamaican Creole (despite a failed petition to make the language official alongside English in 2019 and public debate over appropriate use of the language – exemplified in the polemic responses to a Southwest Airline employee using his native Jamaican Creole in a security announcement, cf. Dawkins 2020). In addition to their lack of consensus about what uses are appropriate for languages in contact, scholars, activists and speakers commonly disagree on the definitions and names of what they are talking about, including whether the mode of communication in question is a new language or a dialectal variety. Although it is not within the scope of this paper to engage with debates of nomenclature regarding contact languages, I hereby acknowledge the complex continuum of codes that exist along a 3-dimensional scale of communication and recognize that for some Puerto Rican bilinguals, such a continuum could be envisioned as having Standard Castilian Spanish at one end and Standard American English at the other, see Figure 1. The central argument of this paper is that the polemic complexity of what lies in between incorporates a Puerto Rican variety of English, which is defined as intelligible to other speakers of English and characterized by its contact with Puerto Rican Spanish, Standard American English, and potentially also Puerto Rican Spanglish, a creole or contact language that may lie at the center of the contact situation.

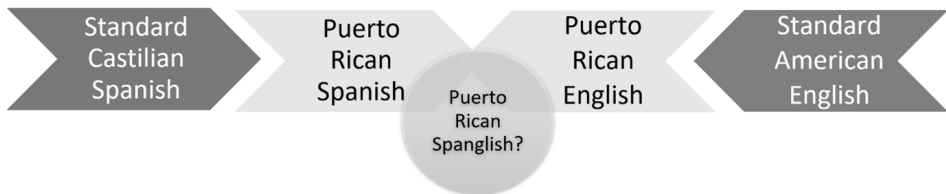


Figure 1: Visual representation of the bilingual language continuum in Puerto Rico

Few studies, with the notable exceptions of Nash (1971) and Walsh (1994), have attempted to document how the English used in Puerto Rico has been affected by contact with Spanish. More recently, Nickels (2005) has advocated for the recognition of Puerto Rican English as a regional dialect that has emerged in a contact situation. The paucity of academic interest in the variety has not made an impact

on public opinion. Most speakers fail to recognize that what they speak is a variety of English; on the other hand, many bilinguals living in Puerto Rico³ downplay the functions and importance of the English they use, which they regularly stigmatize as “inaccurate”, or “deficient.” To compound such speaker perceptions, there are no repositories of the spoken or written English used in Puerto Rico, which speakers might refer to as examples of local usage and academics might use as linguistic data upon which predictions can be formed and tested. In response, this paper tests the hypothesis that speaker involvement in generating a database of linguistic resources on Puerto Rican English can increase awareness of how the Spanish-dominated context of the island has given rise to a distinct variety of English. Specifically, it presents the experimental findings of ongoing work with documentation conducted through a series of pilot studies designed to build capacity at the University of Puerto Rico (UPR) to enable research on mixed language practices and features of Puerto Rican English on the island. The objectives of the work were, firstly, to adapt, devise and pilot-test methodologies that would generate linguistic data for exploratory analysis and, secondly, to establish practices that meaningfully involved participants and early-career scholars in order to help them recognize and value their own Puerto Rican variety of English.

2 The context of the study and related literature

Language mixing practices in Puerto Rico, often referred to with the problematic term “Spanglish”, have a unique contact history in the associated U.S. territory where English and Spanish have been in contact for over a century. Spanish colonialization in the 15th century pre-dates the imposition of English at the turn of the 20th century after the island was ceded to the U.S. after the Spanish-American war. Since a polemic 1902 law designated both English and Spanish as official languages, the status of both languages continues to incite debate, motivate policy change and entrench political divisions (Muñiz-Argüelles 1989). Spanish remains the vernacular in Puerto Rico, but explicit English instruction has long been an objective of the public education system (Torres-González 2002). Despite what Pousada (2006: 17) recognizes as a collective resistance and developmental delay in the acquisition of English in the island, she recognizes that the use of English

³This paper refers to its Puerto Rican study population as bilingual given the educational context of the island; English is taught as a compulsory second language in K-12 grades and is a requirement for two years at the undergraduate level.

and increasing exposure to English media on the island has resulted in “competent bilinguals scattered around the island and concentrated in the coastal and San Juan metro areas” (Pousada 2017: 1). Aligning with this observation, the U. S. Census Bureau (2020: February 17) reports that 23% of the Puerto Rican population speak English “only” or “very well” with the rest of the population reporting proficiency, but less than “very well”.⁴ Thus, for the majority of Puerto Ricans, language contact is perceived as an issue affecting the vernacular Spanish language of the island and little attention has been paid to how it has also affected the English that is spoken on the island.

Aligned with perceptions of Spanglish as a Hispanic issue, much academic interest has focused on how Spanish has been affected by contact with English and on how mixed language practices function among young speakers in Puerto Rico. Moreover, given that the vernacular Spanish operates in diglossia with English for Puerto Rican bilinguals who live on the island, studies conducted in Puerto Rico with bilinguals have invariably focused on Anglicized forms of Spanish or lexical borrowing from English within a Spanish matrix framework (e.g. López Morales 1999, Morales 2000, Holmquist 2013). Thus, academics and speakers alike recognize the existence and characteristics of a Puerto Rican Spanish that is influenced by English at the phonological, morphological, syntactic and discourse levels (see Poplack 1984, Torres 2002, Armstrong 2010, Brown & Rivas 2011). However, there is much less research and much less consensus on the existence and characteristics of a Puerto Rican English that is influenced by Spanish, with most academics passing over the concept completely in favor of considering English use as something intrinsically associated with the mainland U.S., as exemplified in Domínguez-Rosado’s chapter on language issues in the island, entitled “Puerto Rican Spanish and American English” (Domínguez-Rosado 2015: vii). The different historical and social contexts of the two languages, reinforced by their anticipated cultural functions and the prescriptive framework of the Real Academia Española, that all students are taught in school, maintains the division of English and Spanish not only in the educational system and print publication but also in the mental models of speakers. As a result, Puerto Ricans living on the island default to considering the two languages as isolated codes, and so the common practice of mixing the two languages is often stigmatized as inappropriate or incorrect – at best – or considered as indicative of a speaker’s linguistic deficiency – at worst. In addition, the Spanish linguistic heritage of the island that indexes Latino identity is often held in opposition to the English language practices of U.S.

⁴Based on 2018 data indicating a total population of 3,069,357 and 715,370 of those (23.3%) fluent in English.

citizenship, military service, and academic and economic advancement. Hence, language choice and language mixing can indicate (or be perceived to indicate) political orientation and economic status. Yet, among young people, expression of Puerto Rican identity through mediums that permit and even encourage mixed language content, coupled with the scope of language contact in the tourism industry that drives the island's economy and offers many entry-level employment opportunities, has begun to shift perspectives on the acceptability and value of code mixing in professional and social contexts.

As academic interest in language-contact phenomena over recent decades has increased, a growing number of studies on language use in Puerto Rico have explored code-mixing, defined by Odlin (1989) and Muysken (2000) as incorporating a wide range of language contact phenomena including transfer, borrowing, and codeswitching in both language acquisition and fluent bilingual contexts. In an early study on what he called "code-shifting" in Puerto Rico, Lawton suggests that "the constant interplay and interlocking of Puerto Rican Spanish with English suggests incipient creolization" (1979: 257). However, in the four decades since this study was published, no other scholars have argued that the language contact situation has led to an emergent creole. Furthermore, theoretical frameworks for creolization, such as the theory of abnormal transition or nongenetic development (Thomason & Kaufman 1988: 211) do not align with the language contact situation in Puerto Rico where speakers appear to be undergoing a process of contact-induced dialectal divergence in both English and Spanish rather than prompting the creation a new language. Consequently, most recent studies on Puerto Rican bilingualism investigate the influence of context on language choice, for example, the role of social media platforms in computer-mediated communication (Carroll 2008, Carroll & Mari 2017) and the relevance of individual preferences and attitudes that impact codeswitching behaviors among young people (Tamargo & Avilés 2017, Pérez Casas 2016). Such studies on speaker behaviors indicate that Puerto Rican bilinguals increasingly recognize their own language switching practices and are more likely to evaluate such practices favorably (Guzzardo Tamargo et al. 2019). In sum, much of the scholarship on language practices in Puerto Rico focus on either: the local variety of Spanish that is influenced by English, commonly known as Spanglish, or the sociolinguistic context of codemixing practices in informal speech among young people.

English use among Puerto Ricans living on the island was, and still is, a controversial topic (Vélez & Schweers 1993, Dumitrescu 2010, Shenk 2012, González Rivera & Ortiz López 2018). Since the establishment of the free public-school system in Puerto Rico following the U.S. military occupation of the island in 1898, English as a language of instruction has been a recurrent and problematic issue.

Hence, early research on English in Puerto Rico often focused on education (Marvin 1904, Falkner 1908, Desing 1947, Manuel 1952). Towards the second half of the twentieth century, although publications about English use in Puerto Rico continued to have a predominantly pedagogical focus, scholarship also began to address issues of identity and question the imposition of English as a political strategy (Pattee 1945). Publications with a sociolinguistic focus addressed shifting perceptions of English among the Puerto Rican population (Epstein 1968, McCroskey et al. 1985, Clachar 1997) and began to address how growing numbers of English-speaking Puerto Ricans were changing the role of English on the island (Fayer 2000). A handful of scholars have attempted to define the functions and forms of English in Puerto Rico (Nash 1971, Walsh 1994, Nash & Fayer 1996, Fayer et al. 1998), yet there is still little recognition that a local variety of English exists. Nickels' paper on "English in Puerto Rico" (2005) provides an overview of the political, educational, sociolinguistic and literary contexts of the language on the island and explains, "[t]he variety of English spoken in Puerto Rico is only beginning to be identified as a variety in its own right" (234). She additionally recommends that we should start explicitly calling the variety "Puerto Rican English" (Nickels 2005: 235). However, since this study was published, no subsequent scholarship has aimed to explicitly locate Puerto Rican English in a wider discourse on New Englishes. Instead, a handful of recent publications and unpublished dissertations show that there is continued interest in researching the sociolinguistic context of Puerto Rican English as it relates to political and social identity (Falcon 2004, Pérez Casas 2008), the role of English in daily life (Ruíz Correa 2019), and the use of English in Puerto Rican literature (Méndez Alberich 2006). However, the educational importance of English acquisition is still very much at the forefront of scholarship (Ebsworth et al. 2018), and that has taken on a new importance after the economic migration of Puerto Ricans to the mainland U.S. in the aftermath of Hurricane Maria. Yet, very little scholarship, published or otherwise, has offered a detailed description of Puerto Rican English. And only one scholar, Rose Nash, has attempted a comprehensive description of how Puerto Rican English exists on a continuum of language contact in the island; Nash published a series of three papers documenting language contact on the island in terms of English-influenced Spanish (Nash 1970), Spanish-influenced English (Nash 1971) and a type of contact code occupying the middle space (Nash 1982).⁵ However, since this work was published nearly four decades ago, little progress has been made in documenting Puerto Rican English.

⁵Nash includes an endnote in the last of these three publications to explain her intention to write another paper on the second-language Spanish variety of native English speakers on the island. However, if such a paper exists, it remains unpublished.

3 Documentation methodologies

In order to test the hypothesis that speaker involvement in generating a database of linguistic resources on Puerto Rican English can increase awareness of how the Spanish-dominated context of the island has given rise to a distinct variety of English, this experimental study adapted, devised and pilot-tested eight distinct data-collection methodologies relating to spoken language, written language, and surveys on language issues, see Figure 2. All the data-collection activities were conducted in Puerto Rico at the UPR Cayey, a four-year public institution with an enrollment of around 3,000 Hispanic students located in a mountain municipality of the central-eastern region of the island. In all these data-collection activities, undergraduate students of the UPR Cayey were involved as participants and as project-documentation assistants.

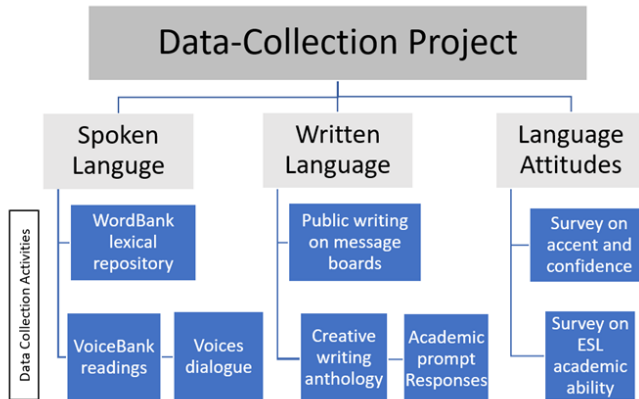


Figure 2: Summary of the eight data collection activities in the pilot project.

3.1 Spoken language

Three distinct data-collection methodologies to capture spoken language were adapted from models devised by the British Library and catalogued in their Sounds archive: 1) a lexical repository called the WordBank, 2) a collection of readings called the VoiceBank, and 3) a series of recorded spontaneous dialogues called Voices. The catalogued holdings of the British Library's ongoing efforts to capture variation in the accents and dialects of English are available online at <https://sounds.bl.uk/Accents-and-dialects/> (British Library Board, n.d.). Using

this website, and specifically the pages linking to the WordBank, VoiceBank and Voices collections, audio files were accessed to serve as a reference point to facilitate the design of the adapted methodology. In addition, descriptions of the methodologies and their associated documentation (including template consent forms) was shared by the lead curator of the British Library's Spoken English archive, Jonathan Robinson, initially in a conference paper (Robinson 2018) and subsequently in meetings and email-communications. As this paper presents the exploratory analysis of only one of these methodologies, the Puerto Rican WordBank, this is the only methodology explained here in detail.⁶

The PR WordBank methodology aimed to record lexical items and their associated meaning, etymology, and usage from among Puerto Rican participants. In the first pilot test of this methodology, conducted between October and December of 2018, instructions given to participants were to think of and then explain a word or phrase in English that they often used, for example with family or friends. Participants were asked to state and spell the word, explain what it means and where it comes from. They were additionally asked to give some context or an example in terms of how they used the word or phrase provided. Recordings were made and catalogued with non-identifying codes that were then associated with the participant's socio-demographic data that they provided on the consent forms. In the second pilot test of this methodology, conducted between August and December of 2019, fieldworkers who were making the recordings similarly instructed participants to say and explain a word or phrase in English that they often used. However, if participants requested more instruction or if they struggled to think of words they might contribute (as happened in the 2018 testing stage), fieldworkers also offered a series of common suggested words from which the participant might choose. This list of possible words and phrases was generated by project-documentation assistants with input from previous work on lexical variation (Nash 1971, Nash & Fayer 1996) supplemented by suggestions from the Project Director and the faculty of the English department of the UPR Cayey based on their experience with lexical variation in local uses of English. The PR WordBank methodology has, to date, generated 142 audio recordings of lexical variation with transcripts and associated metadata. Ongoing exploratory analysis on this data uses techniques in trend analysis to determine patterns in the data which are then analyzed in terms of linguistic parsing and associations with socio-demographic markers.

⁶Interested readers may contact the author at sallyj.delgado@upr.edu for detailed descriptions of any of the methodologies not described in full in this paper.

3.2 Written language

Three distinct data-collection methodologies to document written language were designed to capture variation in both informal and formal contexts for expressive, analytical and academic purposes: 1) public writing on message boards, 2) creative writing for an anthology of narrative work, and 3) developed responses to academic prompts. All written material was collected on location at the UPR Cayey and most participants were undergraduate students of the campus.⁷ As this paper presents the exploratory analysis of only one of these methodologies, viz. public writing on message boards, this is the only methodology explained here in detail.

After the academic calendar resumed on October 23, 2017 following an interruption after Hurricane Maria that made landfall in Puerto Rico on September 20, 2017, a series of public message-boards were posted around campus to spread positive messages in response to written prompts. The prompts, location and dates that message-boards were posted are summarized in Table 1. An example of the photographic images that generated the data is shown in Figure 3. Project-documentation assistants collected the handwritten comments posted on these public message boards with photographic images, re-wrote the comments in electronic documents and determined a system of classification to conduct exploratory research. This methodology to capture public writing on message boards has generated a total of 443 language items, defined as units of meaning as they occur in single words, phrases, or sentences of either one or multiple clauses. Exploratory data-analysis was conducted at three levels. First, a language choice analysis determined participant preferences for either Spanish, English, an alternative language, or mixed code.⁸ Second, a readability analysis was conducted using the Flesch Kincaid Reading Ease test, the Flesch Kincaid Grade Level indicator, the Gunning Fog index, the Coleman Liau index, the Automated Readability Index (ARI), and the SMOG index using the software at https://www.online-utility.org/english/readability_test_and_improve.jsp and at <http://www.readabilityformulas.com/freetests/six-readability-formulas.php>. Lastly, a content analysis was conducted by classifying the focus of responses into four broad categories determined by trends in the data. The four categories related to:

⁷The public writing component captured anonymous contributions, so the research team cannot guarantee that all participants of this data-collection activity were undergraduate students. It is possible that visitors to the campus, faculty, and/or service and administration personnel also contributed to the message boards.

⁸Images and comments that were not decipherable or were composed of unknown acronyms were discounted. Words such as “pizza” that were potentially English, Spanish (or another language) were classified separately so as not to artificially inflate language category results.

Leisure, Community, Physical Fulfillment, and Spiritual Fulfillment, and within these four broad categories, sub-classifications were also determined, e.g., in the “Leisure” category, sub-classifications included: TV, Music, Social Media and Video Games.

Table 1: Summary of the prompts, location and dates that message-boards were posted at the UPR Cayey

Prompt	Location	Date posted	Date collected
<i>Razones para vivir</i>	Outside cafeteria in Student Centre	Sept. 15, 2017	Nov. 27, 2017
positive comments	In corridor of English Department classrooms	Oct. 23, 2017	Nov. 27, 2017
<i>Soy agradecido por...</i>	Outside cafeteria in Student Centre	Nov. 11, 2017	Nov. 27, 2017

3.3 Language attitudes

Two distinct data-collection methodologies to document participant’s attitudes towards language issues were pilot tested at UPR Cayey: 1) a survey on accent and confidence, and 2) a survey on discipline-specific competency using English as a Second Language. Survey responses to questions and prompt statements were given on a Likert scale to generate qualitative data with additional short-response questions to generate qualitative data. In both surveys, quantitative and qualitative data generated from the data collection activities were concurrently analyzed, descriptively and analytically, following a model suggested by Auerbach & Silverstein (2003: 23–28). Quantitative data were processed using the average, maximum, minimum, standard deviation, and the range of the variable created to evaluate the perception of the participants, and qualitative data were analyzed by preparing the transcription of responses and then using exploratory and inferential analysis methods to identify statistically significant trends. Both data-collection methodologies were adapted from previous surveys that had been subject to experimental analysis and evaluated as suitable methods for data-collection. As this paper presents the exploratory analysis of the methodology related to accent and confidence, this is the only methodology explained here in detail.

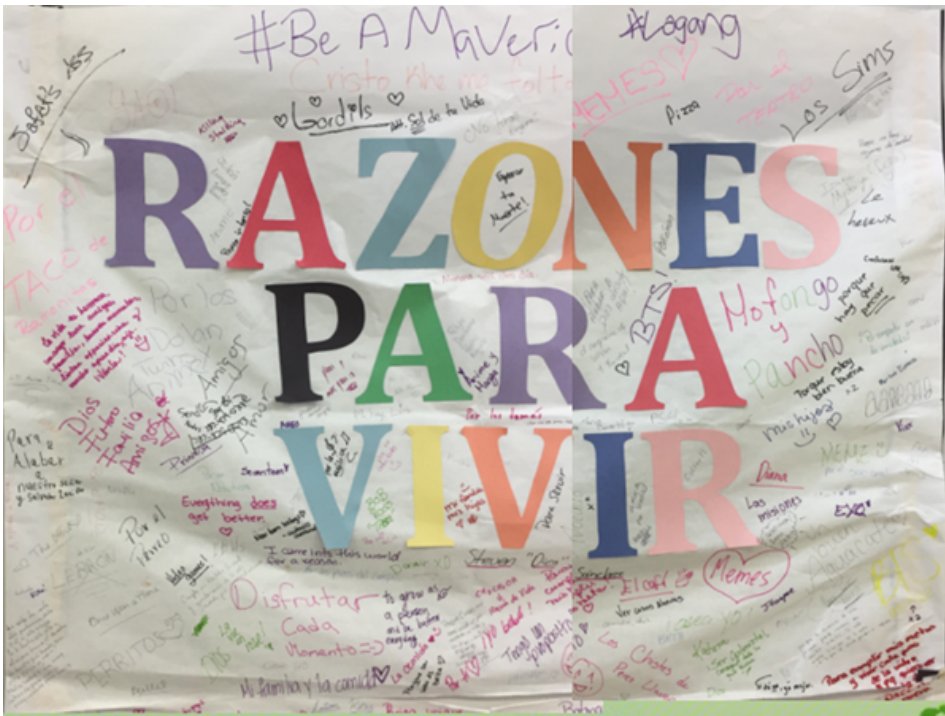


Figure 3: Aligned photographic images taken from the message board “Razones para vivir” that generated data on informal written language.

The UPR Cayey survey on accent and confidence was adapted from a methodology described in a paper entitled “Student’s self-perceived English accent and its impact on their communicative competence and speaking confidence” (Norman 2017). This survey proposes to measure how self-perceptions of accent impact language attitudes and self-confidence in relation to English as a second language in the academic environment, originally among students taking English 6, a course the students must pass to be eligible for further academic studies, in one Upper-Secondary School in Sweden.⁹ The survey adapted for use at UPR Cayey used similar questions as the model questionnaire, that Norman (2017) provided in an index to the published study, with adapted wording that reflected the Puerto Rican context of the new study, for example, the question: “How often do you speak English with someone that is not a native speaker of Swedish?”

⁹Note that this methodology was selected for its experimental design and does not intend to suggest that the language context of Puerto Rico is the same as the Swedish context of the original study.

(Norman 2017: 39) was re-worded to: “How often do you speak English with someone that is not a native speaker of Spanish?” The UPR Cayey study also included demographic information and asked students to answer some questions in a short response, for example: “Is it important to sound like a native speaker of English?” and “Do you think people judge you because of your accent?” The survey was made available from August 16, 2019 and initially promoted among students enrolled in INGL 3265 “English Across Cultures”, before being subsequently shared among a wider community of undergraduates at the UPR Cayey using social media and direct email communication with potential participants. The survey on accent and confidence gathered data from 215 anonymous participants between August 16 and December 29, 2019, of whom the majority self-reported an age of between 16 and 25 and are assumed to be students of the university. The survey was administered using Google Forms, a web-based survey platform that provides rudimentary graphics indicating percentage responses to each question. These graphics and the associated summative data were used to enable exploratory analysis.

4 Results: Exploratory analysis

This section outlines some of the main findings of ongoing exploratory analysis conducted by this author and collaborative groups of early-career scholars at UPR Cayey as they relate to the methodologies of the three data-collection areas of spoken language, written language, and language attitudes.

4.1 Spoken language: PR WordBank

Exploratory research on spoken language using data generated in the PR WordBank indicate three significant categories of local usage that participants recognized as lexical items of Puerto Rican English in addition to usage that might be considered widespread across the U.S. and in African American Vernacular English. The first, and by far the most sizable category representing more than half of the overall contributions at 55%, is composed of English lexicon with comparable meaning to Standard English but demonstrates Spanish bound morphology that typically determines grammatical function. Furthermore, the majority of these lexical items are verbal and have been assigned the default (most common and regular) *-ar* verbal suffix in Spanish.¹⁰ For example, *textear* (‘to text’), *chil-*

¹⁰The orthography of all lexical items indicated in quotation marks derive from participant contributions and/or the determination of Puerto Rican documentation assistants and thus reflect local spelling conventions.

lear ('to chill [out]'), and *postear* ('to post[upload] social media'). Although the majority of the data from this category shows words expressed in the infinitive form, a significant groups of words were also expressed in adjectival participle forms with the affixation of the *-ado* or *-iado* verbal suffix in Spanish that is typically realized as a falling triphthong which is fronted at the end [iao] in Puerto Rican Spanish, for example, *hookiao* ('hooked') which is also orthographically represented as *juqueao* and *jukiao*. Unsurprisingly, because these words have undergone a lexicalization process and are realized in a Spanish matrix framework, their orthography aligns with Spanish norms, as shown in other words such as *chonquear* ('to chunk[vomit]') and *janguear* ('to hang [out]'). Furthermore, phonemes are added if the lexeme defies underlying Spanish phonological rules, e.g., the affixed vocal /e/ that forces the /s/ into the coda of a newly created syllable at the start of the words *esnackear* ('to snack') and *estartear* ('to start'). This category of data reflects a type of interference that (Thomason & Kaufman 1988: 212) describe as resulting from borrowing in the language contact situation and leading to lexical diffusion.¹¹

The second most significant category of the data generated in the PR Word-Bank, representing about a quarter of the overall contributions at 24%, is composed of English-derived words that are adapted for Puerto Rican Spanish phonology and subsequently assigned an extended or alternative semantic scope. Some of these words are derived from English brand names and have undergone extension such as *Conflei* ('Cornflakes') used as a common noun meaning all breakfast cereal, *chicle* ('Chiclets') used as a common noun meaning gum, and *pampel* ('Pampers') used as a common noun meaning diapers. Other words have undergone semantic shift such as "ready", which one participant explains:

So, between my friends and I, and I think most uh part of the island, we use the word "ready" which actually means uh that something is pre-is like prepared. Like if you're ready to go out that means you're prepared to go out, you feel dressed and all that. Um, but we use the word "ready" as like as is a synonym of cool. So, we say, "oh eso está ready", (that's-that's ready, that's cool). It really doesn't have any correlation with it, that's how we use it. (Transcript BL.WB.2018.10.27.KV.IC.1.)

¹¹One caveat is that, although speakers have contributed these terms as examples of their local use of English, they may be considered as features of interference in Puerto Rican Spanish. However, given that these findings are exploratory, and that the conceptual space in which Puerto Rican English manifests necessarily intersects with Puerto Rican Spanish, they are documented here as their speakers intended, as examples of PR English.

Another participant explains how the word “pitcher” is used locally to denote “a person that ignores you or ignores other people” (Transcript BL.WB.2018.11.1.F R.FR.2.). The same participant explains, “It comes from, like, baseball, from pitcher, but I don’t really know what relevance it has with ignoring people but that’s what it means to people here” (Ibid.).

The third, and perhaps the most innovative category of the data generated in the PR WordBank, composing 15% of the overall contributions, is composed of newly coined words or phrases that have a very precise local meaning, for example the term “island dog” that one participant explains is a term for mixed breed stray dogs and is used across the Anglophone Caribbean. Many of these terms that index relationships or salutations are restricted to family or close social networks in usage, and thus may indicate incipient coinage that may yet become widespread, for example, the term *broki* that is described as a synonym for friend and may derive from the English word ‘brother’ or ‘bro’, the word *obb* which one participant describes as a salutation, and the word *peje* that means a close friend or child, and may derive from Spanish *pexe* meaning ‘fish’ or *peje* meaning ‘a cunning person’ or ‘mermaid’ (Roberts 2014: 329). One particularly interesting coinage that was contributed is the term *cangriman* that was explained by one participant as a Puerto Rican adaptation of standard English word *congressman* which his grandfather used to refer to any person with high social status. However, this word also has roots in the Spanish word *cangrejo* (‘crab’) historically used as a derogatory term for foreigners in Latin America and may have been used as a reference to users of cannabis in the late 1970s and 80s because of the way people held their fingers like a crab’s claws to smoke cannabis joints (Urban Dictionary 2007). Hence, the lexicalized expression may have more localized and covert meanings that the contributing morphemes suggest.

4.2 Written language: Public writing on message boards

Exploratory research on public writing using data sourced from public message boards indicate that the language used by participants is strongly influenced by the language used to contextualize the writing activity. Of the three message boards that were used to generate data, the two that used a Spanish prompt: *Razones para vivir* (‘reasons to live’) and *Soy agradecido por...* (‘I am grateful for...’) generated most responses in Spanish (65.5% of responses and 78.8% of responses, respectively). The one message board that used an English prompt that read “The English Department wants to share your positive comments” (subsequently shortened to “Positive comments”) generated most of its responses in English (76.2%), see Table 2. The aggregate data indicate an average of 73.5% of

responses were provided in the same language as the prompt. A greater percentage of responses used code mixing in response to the English prompt rather than the two Spanish prompts, potentially indicating preference for an increased use of the Spanish vernacular in a second language context. In addition, the Spanish prompts generated responses in other languages (e.g., French, Japanese) which was not evident in response to the English prompt, potentially suggesting an increased level of comfort integrating other languages into a Spanish matrix context rather than an English matrix context. It is likely that in addition to the explicit language of the prompt, the location of the message board also implicitly determined the language of responses: the two Spanish-language message boards were located in the student center, a social area where students are more likely to use their Spanish vernacular. The message board with the English prompt was located in the corridor adjacent to English Department classrooms, where students are more likely to use English for academic purposes. Interestingly, this finding appears to corroborate the data in a study of Facebook posts that indicates Puerto Ricans only use English to reply to previous messages written in English (Carroll & Mari 2017).

Table 2: Summary of data on message-boards classified by language of response. *: Same language as the prompt.

prompt	language	language of responses (%)					total
		Spanish	English	mixed	other	misc.	
<i>Razones para vivir</i>	Spanish	65.1*	18.1	2.9	1.7	12.2	100
<i>Soy agradecido por...</i>	Spanish	78.8*	11.8	1.2	1.2	7.0	100
positive comments	English	19.0	76.2*	4.8	0.0	0.0	100

The composition of responses, determined from the application of readability indicators, show that there were a significantly higher percentage of clauses in the message board with the English prompt compared to the two with the Spanish prompts, see Table 3. These results may corroborate the previous finding that responses follow the context of the prompt because neither of the two prompts written in Spanish were clauses, but the English prompt was expressed as a clause. Therefore, just as participants overwhelmingly matched the language of their response to the language of the prompt, participants also matched the grammatical composition of their response to the composition of the prompt. In addition, it is possible that the grid-style organization of the message board with the English prompt may have promoted more clauses as participants may have wanted to

fill a pre-determined space. Comparatively, the organization of the two message boards with the Spanish prompts were presented on a blank background with no suggestion of pre-determined response space, see Figure 3. In addition to phrase and clause analysis, it was determined that the message board with the English prompt had slightly fewer words per clause than the two message boards that used a Spanish prompt (based only on responses that were composed in clauses, as indicated in Table 3, column 4). The message board with the highest number of words per clause was the one with the prompt *Soy agradecido por...* which may have been motivated by the increased formality attached to some of the religious content in responses. There was no significant difference in the average words per syllable among the three message boards.

Table 3: Summary of data on message-boards classified by composition of responses

prompt	prompt composition	percent phrases	percent clauses	average syllables per word	average words per clause
Razones para vivir	phrase	91.6*	8.4	1.5	7.8
Soy agradecido por...	phrase	85.9*	14.1	1.6	8.2
positive comments	clause	8.6	91.4*	1.6	6.5

*Same composition as the prompt

The subject-matter content of responses, determined by the classification of responses into broad categories, showed that participants referred to spiritual fulfillment and community subjects more than leisure or physical fulfillment, see Table 4. This may, in turn, suggest that spiritual fulfillment and a sense of community are more valued than leisure activities or physical fulfillment. Inherent values reflect the nature of the prompts that were designed to focus on subjects of high value to participants that they considered were worth living for, were thankful for, or were positive elements in their lives. This was an important finding for early-career scholars who had initially been shocked by some of the comments related to physical fulfillment, specifically in the sub-categories of sexual desire, (fast)food and alcohol consumption, that had promoted them to initially consider participants' responses as fairly shallow and superficial. After analysis, these same scholars recognized that the number of responses focused on spiritual fulfillment, specifically in the sub-categories of positive self-affirmation and religion, were more than double those of physical needs and suggested that most participants did not respond with superficial comments. Instead, they determined that participants responded with an honesty that indexed both profound

psychological needs and more immediate needs for physical fulfillment and distraction.¹²

Table 4: Summary of data on message-boards classified by content of responses

prompt	content focus of responses (%)				total %
	leisure	physical fulfillment	spiritual fulfillment	community	
Razones para vivir	22.4	18.7	31.1	27.8	100
Soy agradecido por...	18.3	14.0	41.9	25.8	100
positive comments	4.5	8.9	65.7	20.9	100
average percent	15.1	13.9	28.9	24.8	

4.3 Language attitudes: Accent and confidence

Exploratory research on both quantitative and qualitative data collected from the accent and confidence survey shows that Puerto Ricans, who predominantly learn English as a second language, have positive self-perceptions of communicative competence and speaking confidence. Furthermore, the Puerto Rican results are similar to the Swedish participants who use English as a second language in Norman's survey (2017) which the methodology of the present study replicated.¹³ Findings from the Puerto Rican data align with Norman's conclusion from the original study of 80 Swedish participants:

students seem to think that having a native-like accent is overvalued and that communication is to favour over their perceived English accent... most of the students value communication over perceived accent, and many of them say that they do not care how they sound as long as what they say is conveyed. (Norman 2017: ii)

Among the 215 Puerto Ricans who participated in the Puerto Rican survey, a significant majority, representing nearly four out of five participants, said that

¹²The timeframe of this study is notable in that message boards were active both before and after the prolonged academic recess caused by Hurricane Maria, a time in which many Puerto Ricans were experiencing significant trauma, insecurity and anxiety.

¹³Although the history of English in Sweden is not the same as Puerto Rico, a comparable situation of diglossia exists in which 1) advanced competencies in English are required for many programs in institutions of tertiary education, and 2) English competency is a characteristic in elite professions among the middle and upper classes (Berg et al. 2001).

their English is either “very good” or “quite good” and most others evaluated their competency as “average”, see Table 5. Only 8 participants evaluated their English as “not good”, and it is interesting to note that this confidence is notably in excess of the Swedish participants, more than a fifth of whom considered their English “not good” (Norman 2017: 13). Similar to the Swedish participants, when asked if they aim to speak with a specific accent, three out of every five of the Puerto Rican participants said that they do not aim to sound a certain way. Of those students who do aim to speak with an accent, the most common response was a preference for an American accent, followed closely by a preference for a Puerto Rican or Hispanic accent, see Table 6. For those who consciously chose an accent, when asked to indicate the factors that may have influenced their choice, about half of all participants indicated that movies, TV series or online games and social media sites such as Facebook, X (formerly Twitter), Instagram and YouTube have influenced them.¹⁴ However, one in four participants also chose the option “I think it sounds like me.” When explaining their conscious use of a native-like accent (or not) in written responses, many participants recognized intelligibility in addition to educational and social expectation as reasons to sound like a native speaker, for example comments included: “a lot of Americans can’t understand you perfectly if you have a heavy accent”, and “some native english speakers tend to make fun of hispanic accents when speaking English”.¹⁵ Others recognized but did not internalize external pressure, for example: “I believe it is expected, not necessarily important”, and “I dobt[don’t] think it is important, but socially there is a lot of pressure.” However, many responses also celebrated diversity and stressed intelligibility as the main objective of communication, for example, one participant said: “I do not think it is important to sound like a native speaker. I just think it is important that you speak a way the other person is able to understand.” Another commented: “diversity in accents exists so there is no reason to try to sound in another way that isn’t natural to yourself.” Some participants even recognized diversity as a critical component of language change, for example: “Accents are like the spices that allow the food that is the language to expand and diversify. Without accents, language would have a harder time to expand and grow.” Overall, Puerto Rican participants had overwhelmingly positive perceptions of local phonological variation despite their recognition of external pressure to adopt native-like American accents.

¹⁴ Although these social media sites began as predominantly image and text-based platforms, the increasing availability of audio-visual media mean that they are becoming relevant phonological sources for ESL learners.

¹⁵ All syntax, word choice and orthography represent what the participants wrote in their comments.

Table 5: Self-evaluation of competency in English expressed as a percentage of participant responses ($n = 215$)

My English is...	...very good	...quite good	...average	...not good
responses (%)	41.9	36.3	18.1	3.7

Table 6: Participant responses to the question “when you speak English, do you aim to speak with an accent?” expressed as a percentage ($n = 215$)

accent choice	I do not aim to use an accent		I consciously choose an accent	
	American	Puerto Rican	Hispanic or “other”	
accent type responses (%)	63.7	18.6	13.0	4.7

When participants were asked to rank whether it was more important to sound like a native speaker or to express themselves easily, four out of every five participants asserted that ease of expression was more important; fewer participants reported that they had confidence with their accents. In response to a question to determine if participants feel awkward or embarrassed about their accents when they speak English, four out of every five participants said that they feel little-to-no negative feeling, and of the remaining participants, most reported feeling only a little awkward or embarrassed, see Table 7. Participants demonstrated insightful awareness of language variety and high self-confidence in written responses to the question: do you think that a person with English as their first language will understand you better if you sound like a native speaker? Some speakers shifted the burden of comprehension to the listener with comments such as: “Don’t care if they do. While they only speak one language, I speak two. They should make the effort to understand THEIR language regardless the accent” (emphasis by the participant) and “They would need to pay more attention, but if I speak correctly they could still understand me.” Others recognized the scope of variety and varying degrees of intelligibility as natural phenomenon among speakers of any language, for example comments such as “sometimes native speakers don’t understand others very well due to their accents”, “a lot of people aren’t used to hearing other accents” and “It depends on the English accent that they are accustomed[to].” One participant spoke for many with the comment “one should always aim to be understood. But i don’t feel it’s important to hide my accent or pretend my first language is english.” In agreement with such sentiments, it is not surprising that nine in ten participants said they would not

be concerned if they were identified as second language speakers of English, and the two highest factors that participants identified as important were being understood (ranked first) and sounding confident (ranked second), both considered more important in speaking than grammatical accuracy (ranked third).

Table 7: Participant responses to the question “do you feel awkward or embarrassed about your accent when you speak English?” expressed as a percentage ($n = 215$)

Awkward or embarrassed extent of feeling responses (%)	no		yes	
	not at all	not really	a little	often / a lot
	48.4	16.7	27.9	7.0

5 Discussion: Increasing recognition and value of Puerto Rican speech

This section demonstrates how language documentation and exploratory research has increased awareness of the characteristics of Puerto Rican speech, and specifically the existence of Puerto Rican English, among speakers and early-career scholars. It furthermore validates the claim that the methodologies used in this pilot study might be effectively replicated on a larger scale to promote public recognition of a local variety of English and combat some of the social stigma associated with ubiquitous language mixing.

Before working on language documentation activities, responses from undergraduate students generally indicate that many are unaware of dialect varieties, including their own. When asked to define the term “Puerto Rican English”, responses included comments such as: “well, that actually sounds racist”, and “it just means USA English with another accent. There’s nothing deep to it.” Other comments demonstrated misconceptions of language homogeneity: “English is just a language a simple one, there isn’t different types of it.” Unsurprisingly, given the political and economic status of the island, participants generally agreed on the advantages of knowing and being able to use English, as indicated in one comment that “someone bilingual has more open doors.” Yet, students intuitively know that English and Spanish are polarized with each code operating in isolation according the ideologies of exonormative standardization. Students recognized that Spanish is governed by the prescriptive norms of the Real Academia Española of Madrid, Spain, and English is governed by descriptive

U.S. standards that are promoted through the curricula of Puerto Rico's Department of Education and codified in publications such as the Merriam-Webster dictionary. Language mixing is therefore perceived as a dangerous unregulated middle zone that might positively index Latinx identity, particularly in the context of U.S. migration, but can also create uncomfortable points of contrast for Puerto Ricans who live on the island and do not identify with a wider Hispanic identity or (im)migrant status that is often stigmatized in public discourses of the U.S.. Underlying judgements about identity construction through language choices reinforce what Pérez Casas identifies as a characteristic among Puerto Ricans to think of English and Spanish individually despite the common mixing of both languages in daily communication, which she observes "does not necessarily mean that they embrace a 'bicultural' identity" (Pérez Casas 2016: 56–57). Unsurprisingly, given the dominance of Spanish vernacular on the island, the unsettling effects of language contact are typically seen in terms of a Spanish matrix framework, as indicated in one speaker's observation: "the more someone knows English, the more it's integrated into their Spanish." In such ways, most students, before working with the documentation project, demonstrated perceptions that English in Puerto Rico is an encroachment of modern U.S. influence on the Hispanic identity of Puerto Ricans. Consequently, no participants recognized a local variety of English or engaged with ideas about emerging localized norms even when they recognized extensive bilingualism and the profusion of local forms that derive from contact with Spanish.

When asked to explain how Puerto Ricans use language, most students recognize that language mixing is integral to expressions of Puerto Rican identity. Student perspectives on codeswitching generally recognized the phenomena as "normal" yet many also described or implied that it is stigmatized as deficient. In such a context, many participants referred to negative emotions attached to codeswitching practices, for example, one student acknowledged, and appeared to apologize for, how she mixes languages: "I use English and Spanish all the time. *I try not to do it on purpose*" (emphasis added). In other cases, students seemed to justify the stigma of codeswitching by explaining how it is triggered by low competency and lapses of memory, for example, one student explained: "often, speaking more than one language can get a bit confusing sometimes and can make you forget." Many students commented on what they received as their own deficient recall: "I forget how to say a word in Spanish" and "I can't seem to think of the right word... this is specially frustrating." Yet, even when not linked to frustration, participants often made comments that suggest they associate codeswitching with carelessness or comfort: "It's just convenient...I just can't remember certain words in Spanish", and "basically, I use the first language that comes to

mind.” Negative perceptions of mixed codes are not unique to the Puerto Rican context. Indeed, scholarship attests to the commonality of derogatory discourse in contexts where mixed languages emerge, exemplified by movements such as “Speak Good English” in Singapore, which was launched in 2000 in Singapore to diminish the influence of Singlish and promote what was considered a more comprehensible standard variety (Lim 2015: 262). Negative mental models associated with codeswitching have been historically linked with the stigmatized Puerto Rican migrant experience; Tato Laviera, a foundational figure of the Nuyorican Poetry Movement, expresses:

hablo lo inglés matao	[I speak broken (lit. “killed”) English]
hablo lo español matao	[I speak broken (lit. “killed”) Spanish]
no sé leer ninguno bien	[I don’t know how to read either well]

(Lavieria 2014: 7)

Negative mental models are also seemingly legitimized by some of the scholarship, for example Zentella describes a type of codeswitching in Puerto Rico as “crutching” which “occurs when speakers are at a temporary loss for a word or construction in the language which they were speaking before the switch” (Zentella 1982: 49). Two sub-classifications of such crutching, according to Zentella, are “Not Known” when speakers *do not know* the target construction and “Lapse” when speakers *do not recall* the target construction (50).¹⁶ Given the negative associations and wording of scholarship, reinforced by an education system that promotes the separation of the two languages, it is perhaps not surprising to see how young people have adopted similar negative terms to explain their own language mixing choices, which they recognize as part of a more widespread “problem.” This is illustrated by comments in our data such as: “We just forget the word in english and say it in Spanish or vice versa”, and “we don’t remember... in the end we just correct what we are trying to say.” Although this paper does not suggest that the linguistic context for English among bilinguals in Puerto Rico is the same as regions where creoles are spoken, negative attitudes towards Puerto Rican English are comparable to attitudes among speakers of non-standard or non-official languages such as indigenous creoles. For example, García León (2013: 50) found that, in comparison to the positive attitudes towards both Standard English and Standard Spanish, speakers of Trinidadian Creole considered that their Caribbean language was not apt for education or economic

¹⁶I acknowledge that the intention of this author, and many others who codify types of codeswitching, is not to stigmatize speakers but to identify linguistic triggers. However, the framework for discussion often implements a monolingual ideology by assuming that the speaker’s intent is to remain in one language throughout the speech act, when there is often little evidence to support such an assumption.

advancement although it was found to be a strong indicator of national and ethnic identity. Similarly, Puerto Ricans recognize that their local speech practices represent their identity but do not index intelligence or wealth. Consequently, positive orientation toward codeswitching as a marker of local identity appears to force a stigma of deficiency on speakers that, for some, becomes an associated marker of Puerto Rican identity.

However, despite perceptions of individual or collective anxiety about language mixing, after participating in research, many students were able to associate more positive emotions with widespread Puerto Rican codeswitching practices. The same participant who had previously stated “I use English and Spanish all the time. I try not to do it on purpose” later explained how using both languages could be beneficial: “I like to take advantage of both languages... my brain is thinking and processing information in two languages.” And this comment reflects the increasing awareness among education scholars, especially in multilingual contexts around Europe, that translanguaging “promotes a deeper and fuller understanding of the subject matter” (Baker 2001: 281). The word choices of other comments indicated an increased positive association of language mixing, for example, describing it as “our original ‘mezcla’” and explaining local speech practices as “*natural* code switching between English and Spanish” (emphasis added). Students made comments that show how language mixing was not a coping strategy or “crutch” but an integrated cognitive process, as illustrated by one student (the same participant who previously stated “I forget how to say a word in Spanish”) who commented: “My mind *unconsciously combines* both languages to create a message *without separating* both Spanish and English” (emphasis added). Furthermore, the social effects of such recognition had already become apparent to one participant who described a situation when a friend mixed codes creatively and committed what would be perceived as errors in Standard Spanish: “its funny because at first we, his friends, corrected him... but now we noticed that this is his way of speaking.” This recognition aligns with the observations of Pérez Casas that “speakers define and co-construct different identities through a CS[*codeswitching*] style...[and] Their bilingual speech style displays identities that reflect the reality of the two linguistic worlds that coexist in their habitus” (Pérez Casas 2016: 58). Overall, participants recognized and attached positive value to what they recognized as “a new branch into variations of the Spanish/English language” and “a mix of social and culture interactions.”

After working on the documentation activities of the project, when asked what the phrase “Puerto Rican English” means, all participants gave responses had positive value associated with the term and many explicitly recognized the variety as a local dialect. One participant explained, “Puerto Rican English is a dialect

variation of the English language spoken in Puerto Rico. It includes verbal particularities such as phonological or lexical features, and also non-verbal features, like body language or facial expressions.” Others recognized – and rejected – the stigma attached to the local variety of English, for example: “I believe that Puerto Rican English has and is becoming an [sic] important for the identity for many Puerto Ricans, and although it may [be] seen as an incorrect way is [of] speaking English it’s not.” After documenting the effect of language contact on the local variety of English, many participants positively associated Puerto Rican identity with the effects of language contact; one participant explained, “Puerto Rican English to me is a variation of English that is influenced by Puerto Rican Spanish” and another explained, “this phrase [Puerto Rican English] means the combination of spanish and English.” Many participants recognized language mixing as an integral part of Puerto Rican English and one participant described codeswitching as “something that gives us pride.” Most students, after participating in documentation activities, were able to discuss Puerto Rican English in terms that positively indexed local identity rather than U.S. political or economic dependence, for example, one student explained how “the unique puerto rican people mold the English language to there [their] native culture and life style.” Many students were able to recognize non-standard features in positive ways, for example, one student explained, “it means giving another language the puerto rican flare and identity, which makes us stand out from the crowd” and another recognized, “Puerto Ricans have develop *their own* features” (emphasis added). Furthermore, students were able to appreciate these non-standard features in terms of linguistic anthropology, “It’s a dialect with distinctive features that evolved through historical points and represents our culture.” Not only were participants able to positively re-evaluate their own language practices, they demonstrated attitudes indicative of socio-linguistic activism, for example, one participant commented: “The English dialect used by the Puertorican people, one that is growing everyday and *must be accepted* as a dialect of English like any other” (emphasis added). In sum, trend analysis demonstrates overwhelmingly that early-career scholars who participated in the documentation project demonstrated not only increased awareness of the characteristics of Puerto Rican speech but also increased recognition of Puerto Rican English as a local dialect that positively indexes the bilingual identity of its speakers. Furthermore, this positive effect emerges across the spectrum of bilingual participants, whether they described their competency in terms that would render them as incipient, receptive, or functional bilinguals, or whether they reported their abilities as equivalent to balanced, simultaneous bilinguals with native speaker or native-like competency.

The relevance of these exploratory findings, much like the language contact situation in Puerto Rico itself, does not easily compare with other situations in which new dialects of English have emerged. Yet, this documentation project and the positive effects on speaker perceptions it has engendered may benefit from contrast with studies on attitudes towards varieties of English that have emerged in similar socio-historical situations. For example, educated young Puerto Ricans, much like the Saudi Arabian students of the study by Al-Dosari (2011: 1047), value listener comprehension over perceptions of native fluency or prestige when evaluating the accents of second language speakers. The speakers of Puerto Rican English involved in the data-collection activities of this study appear similar to speakers of Indian English who display positive attitudes towards their own variety of English despite residual preference for the prestige variety associated with colonial history (Bernaisch & Koch 2015). In terms of usage, the complexities of language convergence and divergence between a local variety of English alongside Standard American English and Puerto Rican Spanish that Puerto Ricans experience could be compared to linguistic accommodation in the multilingual context of Singapore (Ng et al. 2014). Tupas (2016) explains how recognizing both the local variety of English in Singapore, in addition to an international standard, can promote attitudinal change that facilitates second language acquisition without compromising cultural identity. Perhaps, as studies and participant involvement in documentation activities continue to promote awareness of the Puerto Rican variety of English, the island might undergo steps towards additive bi-dialectalism in language classrooms that address perpetual issues of “non-standard” usage in such contexts, comparable to concerns in Singapore (Tan & Tan 2008). However, Puerto Rico has a unique language contact situation, unlike the multilingual environments of Singapore and India. Pousada (2006: 17) explains:

Puerto Rico is distinct from other countries such as Singapore, Hong Kong, the Philippines, and India where English has been successfully implanted. Those countries are linguistically very heterogeneous and have acquired a local variety of English for diplomatic, commercial, and technological communication among diverse populations. For them, English is an ethnically neutral language that does not threaten their nationality and is utilized primarily as a lingua franca for pragmatic purposes. In contrast, in Puerto Rico, because of its historical domination by the United States, planning for improving English learning is often viewed with suspicion as an attempt to unseat Spanish which is the native language of almost all residents on the island.

Thus, although comparative approaches with other contexts of new Englishes may be informative, they are unlikely to provide satisfactory models for Puerto Rican English, a variety trapped between two colonial giants on a binary continuum. Perhaps, instead of looking towards studies on World Englishes, comparative studies on Hispanic varieties of English in the U.S. might provide better context for the language contact situation in Puerto Rico. In their study of Hispanic English in the mid-Atlantic south, Wolfram et al. (2004) highlight the importance of examining the dynamic early stages of English dialect emergence in contact with a local variety of Spanish. Perhaps studies on Puerto Rican English, with its geographic and cultural distance from mainstream U.S. influence, might provide comparative insights that inform what we know about how Hispanic communities play a role in forming the symbolic role of regional varieties of English among both second language speakers and multilingual communities.

6 Conclusions and recommendations

The data presented in this paper serves to support the hypothesis that speaker involvement in generating a database of linguistic resources on Puerto Rican English can increase awareness of how the Spanish-dominated context of the island has given rise to a distinct variety of English. The eight methodologies that have been tested as part of this pilot project in addition to the experimental analysis that has been conducted on their data demonstrate that they are feasible strategies to begin documenting the characteristics of Puerto Rican English. The scope and meaningful involvement of participants and early-career scholars has demonstrated that there is interest in local variation. Furthermore, the effects of their involvement in documentation activities indicate that participants not only demonstrate increased recognition and positive evaluation of Puerto Rican English, but also recognize language mixing as an integral expression of Puerto Rican identity rather than a marker of linguistic deficiency.

Exploratory research suggests that, without the opportunity to study their own language variation, young Puerto Ricans do not recognize that they speak a unique dialect of English. Furthermore, they continue to attach stigma to their own mixed language practices and often consider mixing languages in terms of how English transfer marks linguistic deficiency in Spanish. They are more likely to perceive code-switching as something only appropriate to informal contexts, thus restricting the potential to use translanguaging practices to maximum advantage in an academic environment. However, young Puerto Ricans also recognize that their mixed language practices are a key component of their bilingual

identity and they typically mix languages in emotive contexts and in ways that bolster social unity and emotional health. In addition, perception surveys indicate that young Puerto Ricans are increasingly confident about their language abilities and reject attitudes that might stigmatize their accents and language mixing practices. Based on the findings of these pilot studies in documenting Puerto Rican English, a central recommendation of this paper encourages speakers, educators, community leaders and other stakeholders in Puerto Rican English to document the variety and advocate for its local and international recognition. Perhaps as we continue to document and discuss Puerto Rican English, more speakers might share the feelings of one participant who attended and gave feedback on a student presentation of exploratory findings:

After listening to this research presentation, I surely felt less insecure, more confident about my “accent.” We spend so much time trying to sound perfect when being understood is enough. I wasn’t really aware that Puerto Rican English was actually a thing, but I’m glad it is and that I know of it.

Acknowledgements

This work is dedicated to the many participants and the documentation assistants of the University of Puerto Rico in Cayey who helped provide authentic language samples and local interpretations about their usage. Without them, none of this would have been possible.

References

- American Heritage Dictionary of the English Language. 2020. “Spanglish”. *Houghton mifflin harcourt*. %7Bhttps://ahdictionary.com/word/search.html?q=Spanglish%7D (2 September, 2019).
- Ardila, Alfredo. 2005. Spanglish: An anglicized Spanish dialect. *Hispanic Journal of Behavioral Sciences* 27(1). 60–81. DOI: 10.1177/0739986304272358.
- Armstrong, Meghan E. 2010. Puerto Rican Spanish intonation. In Pilar Prieto & Paolo Roseano (eds.), *Transcription of intonation of the Spanish language*, 155–189. Munich: LINCOM.
- Auerbach, Carl & Louise B Silverstein. 2003. *Qualitative data: An introduction to coding and analysis*. New York: New York University Press.
- Baker, Colin. 2001. *Foundations of bilingual education and bilingualism*. 3rd edn. Clevedon, England: Multilingual Matters.

- Berg, E. Cathrine, Francis M. Hult & Kendall A. King. 2001. Shaping the climate for language shift? English in Sweden's elite domains. *World Englishes* 20(3). 305–319. DOI: 10.1111/1467-971x.00217.
- Bernaisch, Tobias & Christopher Koch. 2015. Attitudes towards Englishes in India. *World Englishes* 35(1). 118–132. DOI: 10.1111/weng.12174.
- British Library Board. 2018. *Sounds: Accents & dialects [online database of audio recordings]*. <https://sounds.bl.uk/Accents-and-dialects> (8 November, 2018).
- Brown, Esther L. & Javier Rivas. 2011. Subject-verb word order in Spanish interrogatives. *Spanish in Context* 8(1). 23–49. DOI: 10.1075/sic.8.1.02bro.
- Carroll, Kevin S. 2008. Puerto rican language use on myspace. Com. *CENTRO Journal* 20(1). 96–111.
- Carroll, Kevin S & Vanessa Z Mari. 2017. Puerto Rican language use on Facebook. *CENTRO Journal* 29(2). 38–61.
- Clachar, Arlene. 1997. Students' reflections on the social, political, and ideological role of English in Puerto Rico. *Hispanic Journal of Behavioral Sciences* 19(4). 461–478. DOI: 10.1177/07399863970194005.
- Dawkins, Nickesha. 2020. *Social commentary on the acceptance and use of creole languages by international airlines to improve the efficiency of border control security in the United States of America*. Linguistic Society of America (LSA) 94th Annual Meeting, New Orleans, January 2-5, 2020. New Orleans: Linguistic Society of America (LSA).
- Desing, Minerva F. 1947. English at the University of Puerto Rico. *College English* 9(2). 91–98.
- Domínguez-Rosado, Brenda. 2015. *The unlinking of language and Puerto Rican identity: New trends in sight*. Cambridge: Cambridge Scholars Publishing. 117.
- Al-Dosari, Hamad Shabeb. 2011. An investigation of attitudes towards varieties of spoken English in a multi-lingual environment. *Theory and Practice in Language Studies* 1(9). 1041–1050. DOI: 10.4304/tpls.1.9.1041-1050.
- Dumitrescu, Domnita. 2010. Spanglish: An ongoing controversy. In Susana Rivera-Mills & Juan Antonio Trujillo (eds.), *Building communities and making connections*, 136–167. Cambridge Scholars Publishing: Newcastle.
- Ebsworth, Miriam Eisenstein, Timothy John Ebsworth & Chencen Cai. 2018. English acquisition in Puerto Rico: Teachers' insights. *Bilingual Research Journal* 41(1). 69–88. DOI: 10.1080/15235882.2017.1413441.
- Epstein, Erwin H. 1968. Linguistic orientation and changing values in Puerto Rico. *International Journal of Comparative Sociology* 9(1). 61–76. DOI: <https://doi.org/10.1163/156854268X00032>.

- Falcon, Sandra M. 2004. *Attitudes toward languages in contact: Spanish, English, and intellectuals in Puerto Rico*. Retrieved from ProQuest Dissertations & Theses Global. (3161622).
- Falkner, Roland P. 1908. Training of teachers in Porto Rico. *Journal of Education* 67(22). 596–597. DOI: 10.1177/002205740806702205.
- Fayer, Joan, Judith Castro, Marta Díaz & Miriam Plata. 1998. English in Puerto Rico. *English Today* 14(1(53)). 39–44.
- Fayer, Joan M. 2000. Functions of English in Puerto Rico. *International Journal of the Sociology of Language* 142(1). 89–102. DOI: 10.1515/ijsl.2000.142.89.
- García León, Javier E. 2013. Being successful meant speaking well, and speaking well meant speaking standard English”. Aproximación al estudio de las actitudes lingüísticas en hablantes trilingües: Inglés criollo, inglés estándar y español. *Forma y Función* 26(1). 11–53.
- Githiora, Chege. 2002. Sheng: Peer language, Swahili dialect or emerging Creole? *Journal of African Cultural Studies* 15(2). 159–181. DOI: 10.1080/1369681022000042637.
- González Rivera, Melvin & Luis A. Ortiz López. 2018. El español y el inglés en Puerto Rico: Una polémica de más de un siglo. *CENTRO Journal* 30(1). 106–131.
- Guzzardo Tamargo, Rosa E, Verónica Loureiro-Rodríguez, Elif F Acar & Jessica Vélez Avilés. 2019. Attitudes in progress: Puerto Rican youth’s opinions on monolingual and code-switched language varieties. *Journal of Multilingual & Multicultural Development* 40(4). 304–321.
- Hernández, Carmen D. 1997. *Puerto Rican voices in English: Interviews with writers*. Westport: Praeger.
- Holmquist, Jonathan. 2013. Spanish/English contact in rural Puerto Rico. *Spanish in Context* 10(3). 390–409. DOI: 10.1075/sic.10.3.04hol.
- Laviera, Tato. 2014. *Benedición: The complete poetry of tato laviera*. Houston: Arte Público Press.
- Lawton, David. 1979. Code-shifting in puerto rican Spanish/English. *Linguistics* 17(3/4). 257–265.
- Lim, Lisa. 2015. Coming of age, coming full circle: The (re)positioning of (Singapore) English and multilingualism in Singapore at 50. *Asian Englishes* 17(3). 261–270. DOI: 10.1080/13488678.2015.1090772.
- Lipski, John. 2004. *Is “Spanglish” the third language of the South? Truth and fantasy about U. S. Spanish. In Language Variation in the South (LAVIS III) 3rd Conference, Tuscaloosa, AL (Vol. 6)*. <https://lavis.as.ua.edu/handouts/lipski.pdf> (12 May, 2019).

- López Morales, Humberto. 1999. Anglicismos en el léxico disponible de Puerto Rico. In Luis A. Ortiz Lopez (ed.), *El caribe hispánico: Perspectivas lingüísticas actuales, homenaje a manuel álvarez nazario*, 147–170. Madrid: Iberoamericana Vervuert.
- Manuel, Herschel T. 1952. Results of a half-century experiment in teaching a second language. *The Modern Language Journal* 36(2). 74–76. DOI: 10.1111/j.1540-4781.1952.tb06110.x.
- Marvin, A. W. 1904. Education in porto rico. *Education* 24(9). 562–564.
- McCroskey, James C., Joan M. Fayer & Virginia P. Richmond. 1985. Don't speak to me in English: Communication apprehension in Puerto Rico. *Communication Quarterly* 33(3). 185–192. DOI: 10.1080/01463378509369597.
- Méndez Alberich, Johanna. 2006. *English in a traffic jam in "La guaracha del Macho Camacho"*. Retrieved from ProQuest Dissertations & Theses Global. (3236760).
- Montes-Alcalá, Cecilia. 2018. Bilingualism and biculturalism. In Francisco A. Lomelí, Denise A. Segura & Elyette Benjamin-Labarthe (eds.), *Routledge handbook of chicana/o studies*, 316–329. New York: Routledge. DOI: 10.4324/9781315726366-30.
- Morales, Amparo. 2000. ¿Simplificación o interferencia?: El español de Puerto Rico. *International Journal of the Sociology of Language* 142(1). 35–62. DOI: 10.1515/ijsl.2000.142.35.
- Muñiz-Argüelles, Luis. 1989. The status of languages in Puerto Rico. In Paul Pupier & José Woehrling (eds.), *Langue et droit*, 457–472. Montreal: Wilson & Lafleur.
- Muysken, Pieter. 2000. *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.
- Nash, Rose. 1970. Spanglish: Language contact in Puerto Rico. *American Speech* 45(3/4). 233–233.
- Nash, Rose. 1971. Englañol: More language contact in Puerto Rico. *American Speech* 46. 106–122.
- Nash, Rose. 1982. Pringlish: Still more language contact in Puerto Rico. In Braj B Kachru (ed.), *The other tongue: English across cultures*, 250–269. Urbana: University of Illinois Press.
- Nash, Rose & Joan M. Fayer. 1996. Changing lexical features in puerto rican English. *World Englishes* 15(3). 281–294. DOI: 10.1111/j.1467-971x.1996.tb00115.x.
- Ng, Chin Bee, Francesco Cavallaro & Daphine Shu Ping Koh. 2014. Singlish can and speech accommodation in Singapore English. *World Englishes* 33(3). 398–412. DOI: 10.1111/weng.12097.

- Nickels, Edelmira L. 2005. English in Puerto Rico. *World Englishes* 24(2). 227–238. DOI: 10.1111/j.1467-971x.2005.00406.x.
- Norman, Johanna. 2017. *Student’s self-perceived English accent and its impact on their communicative competence and speaking confidence: An empirical study among students taking English 6 in upper-secondary school*. <https://www.divaportal.org/smash/get/diva2:1146855/FULLTEXT01.pdf> (19 December, 2018).
- Odlin, Terence. 1989. *Language transfer, cross-linguistic influence in language learning*. Cambridge: Cambridge University Press. 224.
- Otheguy, Ricardo & Nancy Stern. 2010. On so-called spanglish. *International Journal of Bilingualism* 15(1). 85–100. DOI: 10.1177/1367006910379298.
- Oxford English Dictionary. 1989. *Oxford English dictionary, 2nd edition*. Oxford, England: Clarendon Press.
- Pattee, Richard. 1945. Puerto Rico and the language question. *America* 74(13). 343–345.
- Pérez Casas, Marisol. 2016. Spanish-English codeswitching in the Caribbean and the US. In E. Guzzardo Tamargo, Catherine M Mazak & M Carmen Parafita Couto (eds.), *Spanish-English codeswitching in the Caribbean and the US*, 37–60. Amsterdam & Philadelphia: John Benjamins.
- Pérez Casas, Marisol. 2008. *Codeswitching and identity among island Puerto Rican bilinguals*. Retrieved from ProQuest Dissertations & Theses Global. (3341614).
- Poplack, Shana. 1984. Variable concord and sentential plural marking in puertorican Spanish. *Hispanic Review* 52(2). 205–222.
- Pousada, Alicia. 2006. Fomenting bilingualism through language awareness. *PRTESOLGram* 33(1). 17–19.
- Pousada, Alicia. 2017. *Being bilingual in borinquen: Student voices from the university of Puerto Rico*. Cambridge: Cambridge Scholars Publishing. 192.
- Real Academia Española. 2020. “Espanglish”. <https://dle.rae.es/espanglish#Rq8JJSS> (2 September, 2019).
- Roberts, Edward A. 2014. *A comprehensive etymological dictionary of the Spanish language with families of words based on Indo-European roots*. Bloomington: Xlibris Corporation.
- Robinson, Jonathan. 2018. *Puggling through kintle for Bobby Dazzlers: The evolving English VoiceBank. 5th conference of the International Society for the Linguistics of English (ISLE), University College, London July 18, 2018*.
- Romaine, Suzanne & Braj Kachru. 1992. Code-mixing and code-switching. In Tom McArthur (ed.), *The Oxford companion to the English language*, 228–229. Oxford & New York: Oxford University Press.

- Ruíz Correa, Sandra. 2019. *The evolution of the English language in daily life in Puerto Rico: A socio-historical analysis of recent events and their impac*. Retrieved from ProQuest Dissertations & Theses Global. (13879720).
- Sayer, Peter. 2008. Demystifying language mixing: Spanglish in school. *Journal of Latinos and Education* 7(2). 94–112. DOI: 10.1080/15348430701827030.
- Shenk, Elaine. 2012. Puerto Rico's language officialization debates: Perspectives from a mountain corridor community. *CENTRO Journal* 24(1). 90–119.
- Stavans, Ilon. 2000. Spanglish: Tickling the tongue. *World Literature Today* 73(3). 555–555.
- Tamargo, Rosa E. Guzzardo & Jessica Vélez Avilés. 2017. La alternancia de códigos En Puerto Rico: Preferencias Y actitudes. *Caribbean Studies* 45(1-2). 43–76. DOI: 10.1353/crb.2017.0003.
- Tan, Peter K & Daniel Tan. 2008. Attitudes toward non-standard English in Singapore. *World Englishes* 27(3-4). 465–479. DOI: 10.1111/j.1467-971x.2008.00578.x.
- Thomason, S & T Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Los Angeles: University of California Press. 427.
- Torres, Lourdes. 2002. Bilingual discourse markers in Puerto Rican Spanish. *Language in Society* 31(1). 65–83. DOI: 10.1017/s0047404502001033.
- Torres-González, Roamé. 2002. *Idioma, bilingüismo y nacionalidad: La presencia del inglés en Puerto Rico*. San Juan, Puerto Rico: La Editorial de la Universidad de Puerto Rico. 431.
- Tupas, Ruanni. 2016. Singlish in the classroom: Is Singapore ready for additive bidialectalism? *International Journal of Bilingual Education and Bilingualism* 21(8). 982–993. DOI: 10.1080/13670050.2016.1226757.
- U. S. Census Bureau. 2020. *Language spoken at Home*. https://data.census.gov/cedsci/table?q=0400000US72&tid=ACSST1Y2018.S1601&t=Language%5C%20Spoken%5C%20at%5C%20Home&vintage=2018&hidePreview=true&cid=S1601%5C_C01_001E (15 January, 2020).
- Urban Dictionary. 2007. “*cangriman*” [lexical definition]. <https://www.urbandictionary.com/define.php?term=Cangriman> (12 May, 2018).
- Vélez, Jorge A & C William Schweers. 1993. A U. S. Colony at a linguistic crossroads: The decision to make Spanish the official language of Puerto Rico. *Language Problems and Language Planning* 17(2). 117–139.
- Walsh, Maureen. 1994. *Towards a Puerto Rican variety of English: An ethnographic study with implications for the English curriculum in Puerto Rico*. University of Puerto Rico. (Unpublished EdD dissertation).
- Wolfram, Walt, Phillip Carter & Beckie Moriello. 2004. Emerging Hispanic English: New dialect formation in the American South. *Journal of Sociolinguistics* 8(3). 339–358. DOI: 10.1111/j.1467-9841.2004.00264.x.

- Yumoto, Kazuko. 1995. Bilingualism, code-switching, language mixing, transfer and borrowing: Clarifying terminologies in the literature. *Kanagawa Prefectural College of Foreign Studies, Working Papers* 17. 49–60.
- Zabrodskaia, Anastassia. 2013. Morphosyntactic contact-induced language change among young speakers of Estonian Russian. In Isabelle Léglise & Claudine Chamoreau (eds.), *The interplay of variation and change in contact settings*, 77–106. Amsterdam: John Benjamins.
- Zentella, Ana Celia. 1982. Spanish and English in contact in the United States: The Puerto Rican experience. *Word* 33(1-2). 41–57. DOI: 10.1080/00437956.1982.11435721.

Name index

- Aboh, Enoch Oladé, 141, 146, 147
Agheyisi, Rebecca, 112, 117
Aguirre Licht, Daniel, 146
Ahlbäck, Anders, 30, 31, 40
Al-Dosari, Hamad Shabeb, 203
Alleyne, Mervyn Coleridge, 100
Álvarez López, Laura, 2
Ameka, Felix, 65, 71, 72, 146
Amfo, Nana Aba Appiah, 148
Anchimbe, Eric A, 119
Anderson, Benedict, 113
Ardila, Alfredo, 178
Arends, Jacques, 15, 19
Armstrong, Meghan E, 182
Asmussen, Jesper Bering, 20
Assis, Virgínia Maria Almôedode,
158, 160
Auerbach, Carl, 188
Auerbach, Joh. Christoph, 20
Avilés, Jessica Vélez, 183
- Baker, Colin, 201
Baker, Philip, 21, 100
Bakker, Peter, 20, 43, 47, 98
Ballester, Maria Teresa Galarza, 21
Barac, Vladimir, 15, 25, 31, 37, 38, 40
Barbosa, Maria do Socorro Ferraz,
163
Bartens, Angela, 21, 114, 118, 121
Bello, Mariama, 38
Berg, E. Cathrine, 195
Bernaisch, Tobias, 203
- Berschin, Helmut, 117
Berthele, Raphael, 71
Birmingham, David, 40
Blench, Roger, 23
Blommaert, Jan, 118, 120
Bøegh, Kristoffer Friis, 19, 20, 23, 48
Bouwer, Leoni E, 98, 99
Bowern, Claire, 98
Brousseau, Anne-Marie, 38
Brown, Esther L., 182
Bryan, M. A., 36
Brydon, Diana, 119
Bucheli Berger, Claudia, 73, 74, 85
- Calabrese, Rita, 119
Capo, Hounkpati B. C, 141
Carroll, Kevin S, 183, 193
Cassidy, Frederic Gomes, 114
Chavagne, Jean-Pierre, 166–168
Chernobilsky, Ellina, 98
Clachar, Arlene, 184
Clyne, Michael, 3, 99, 120
Codó, Eva, 116
Coleman, William Donald, 119
Colmenares, Germán, 140
Coutinho, Ismael de Lima, 166
Csizér, Kata, 117
- da Costa Peixoto, António, 21
da Hora, Dermeval, 166
da Silva Sobrinho, José Simão, 99
da Silva, Nilce, 99

- Dalby, David, 39
Davies, Mark, 150
Dawkins, Nickesha, 180
de Josselin de Jong, J.P.B., 46
DeGraff, Michel, 1, 88, 118, 121
Desing, Minerva F, 184
Deuber, Dagmar, 98
Dieck, Marianne, 135
Dollinger, Stefan, 3, 100, 122
Domínguez-Rosado, Brenda, 182
Dörnyei, Zoltán, 117
Doucet, Rachelle Charlier, 117
Dryer, Matthew S., 23, 147
Dumitrescu, Domnita, 179, 183
- Eberhard, David M., 23
Ebsworth, Miriam Eisenstein, 184
Eckert, Penelope, 121
Eisenlohr, Patrick, 121
Ellis, Alfred B, 146
Elspaß, Stephan, 3, 100
Eltis, David, 22, 26, 141
Epstein, Erwin H., 184
Essegbey, James, 65, 71, 72
- Falcon, Sandra M, 184
Falkner, Roland P., 184
Faraclas, Nicholas, 3, 100, 101
Faria, Ernesto, 166
Fattier, Dominique, 73
Fayer, Joan, 184
Fayer, Joan M., 184, 186
Feldbæk, Ole, 20
Ferraz, Luis I, 118, 135
Fishman, Joshua A., 112, 117, 118
Fleischmann, Ulrich, 1
Flórez, Luis, 135
Fodor, István, 21, 31, 35, 37–40
Fraser, Adrian, 21, 23
- Freeland, Jane, 113, 118
- Gaines, Richard, 71
Galves, Charlotte, 158
García León, Javier E, 200
García, Ofelia, 119
Giles, Howard, 120
Githiora, Chege, 179
Gøbel, Erik, 17, 18, 25, 26, 29
Gonçalves, Perpétua, 99
González Rivera, Melvin, 183
Goodman, Morris F, 47
Graebin, Geruzade Souza, 158, 168
Granda, Germán de, 135, 142, 148, 150
Green-Pedersen, Svend E, 17, 18, 29, 43
Güldemann, Tom, 23, 135
Gumperz, John Joseph, 121
Guthrie, Malcolm, 22, 39, 40
Guzzardo Tamargo, Rosa E, 183
- Hagemeijer, Tjerk, 135
Hair, Paul E. H., 25, 39
Hall, Gwendolyn M, 44
Hall, Gwendolyn M., 144
Hall, Neville T, 17–19
Hall, Robert Anderson, 98
Hancock, Ian Francis., 100, 116
Harms, Phillip Lee, 145
Haspelmath, Martin, 23
Haugen, Einar, 119
Hernández, Carmen D, 179
Herskovits, Melville L, 40
Highfield, Arnold R., 15, 19, 23–25, 31, 37, 38, 40, 48
Higman, B. W, 15, 18
Hollington, Andrea, 108
Holm, John, 1, 2, 100
Holmer, Nils M, 146

Name index

- Holmquist, Jonathan, 182
Holsoe, Svend E, 19, 49, 51
Hopkins, Daniel P, 18
Hornby, Ove, 18
Huber, Magnus, 100, 106, 107
Hüning, Matthias, 100, 120
Hvid, Mirjam Louise, 25
- Inikori, Joseph E, 2
Innes, Gordon, 36
Inverno, Liliana C. C, 171
- Jensen, Niklas Thode, 14, 16, 17
Jespersen, Otto, 150
Jiménez Meneses, Orián, 139
Jones, Adam, 21, 31, 35–37, 40
Juruá, Padrinho, 160
Justesen, Ole, 20
- Kachru, Braj, 179
Kaufman, T, 183, 191
Keenan, Elinor., 111
Kelley, Sean M., 37
Klein, Herbert S., 159
Kloss, Heinz, 3, 99, 100
Koch, Christopher, 203
Koelle, Sigismund W, 35, 39
Kopecka, Anetta, 65, 69–71
Koskinen, Arja., 118
Kouwenberg, Silvia, 15, 21, 30
Krämer, Philipp, 100, 120
Kubota, Ryuko, 120
- Laban, Michel, 166, 168
Labov, William, 111, 121
Lakoff, Robin T, 111
Lambert-Brétière, Renée, 65, 71, 72,
93
Larsen, Jens, 24
Laviera, Tato, 200
- Law, Robin, 35, 44, 141, 144
Lawaetz, H, 19
Lawton, David, 183
Le Page, Robert Brock, 114
Leal, Claudia, 139
Lefebvre, Claire, 38
Levi, Joseph Abraham, 158
Levisen, Carsten, 119
Li, Wei, 119
Lim, Lisa, 200
Lipski, John, 178
Lobo, Tânia, 158
Loewen, Jacob A, 139
López Morales, Humberto, 182
Lovejoy, Paul E., 37
Lucchesi, Dante, 157–159, 164, 171
Lumwamu, François, 148
Lüpke, Friederike, 118
- Mac Cord, Marcelo, 158
Magens, J.M., 19
Mair, Christian, 100, 121
Makoni, Sinfree, 100, 119, 120
Manuel, Herschel T., 184
Mari, Vanessa Z, 183, 193
Marvin, A. W., 184
Massini-Cagliari, Gladis, 158
Matras, Yaron, 98
Mattos e Silva, Rosa Virgínia, 158,
164
Maurer, Philippe, 135, 148, 149
McConnell-Ginet, Sally, 121
McCroskey, James C., 184
McWhorter, John, 100
McWhorter, John H, 98, 100, 116, 141
McWhorter, John H., 1
Megenney, William W, 135
Meier, Gudrun, 23, 30
Méndez Alberich, Johanna, 184

- Michaelis, Susanne M., 100
Migeod, Frederick W. H., 35
Moll, Andrea, 108, 121
Montes Giraldo, Joaquin, 135
Montes-Alcalá, Cecilia, 178
Morales, Amparo, 182
Mortenson, Charles A., 145
Mosquera, Sergio A., 144
Moysés, Sarita M. Affonso, 162
Mufwene, Salikoko S., 2, 47, 119, 120
Mühleisen, Susanne, 100, 119, 121
Muhr, Rudolf, 3, 99, 120
Muñiz-Argüelles, Luis, 181
Mutz, Kathrin, 73
Muysken, Pieter, 19, 21, 179, 183
- Nash, Rose, 180, 184, 186
Ng, Chin Bee, 203
Nickels, Edelmira L., 180, 184
Niehaus, Konstatin, 3, 100
Nielsen, Herluf, 26
Noll, Volker, 5, 158, 169
Norman, Johanna, 189, 190, 195, 196
Nørregaard, Georg, 25
- Odlin, Terence, 179, 183
Ogay, Tania, 120
Oldendorp, Christian Georg Andreas, 2, 15, 19–21, 30–32, 35–40, 44
Oliveira, Klebson, 157–159, 164, 166–169
Olsen, Poul Erik, 17, 18, 24, 29, 30, 54
Ortiz López, Luis A., 135, 183
Otheguy, Ricardo, 178
Ottó, Istvan, 117
- Palmié, Stephan, 2
Parkvall, Mikael, 20, 21, 25, 46, 47
- Patrick, Peter, 100
Pattee, Richard, 184
Pennycook, Alastair, 100, 120
Pérez Casas, Marisol, 183, 184, 199, 201
Petter, Margarida, 5, 159, 164, 166–169, 171
Pontoppidan, Erik, 20
Pope, Pauline Holman, 2, 15, 20, 48–53
Poplack, Shana, 182
Post, Marike, 135
Pourcel, Stéphanie, 65, 69–71
Pousada, Alicia, 179, 181, 182, 203
Prescod, Paula, 21, 23
- Quintão, Antonia Aparecida, 158, 160
- Rettich, Margret, 75
Rettich, Rolf, 75
Richardson, David, 141
Rivas, Javier, 182
Roberts, Edward A., 192
Roberts, Justin, 17, 18
Robinson, Jonathan, 186
Romaine, Suzanne, 98, 111, 179
Ruíz Correa, Sandra, 184
Ruíz-García, Martha, 4, 134–136, 139, 150
- Sabino, Robin, 19, 20, 25, 27, 46, 47
Sayer, Peter, 179
Schaefer, Ronald P., 71
Scheuringer, Hermann, 99
Schieffelin, Bambi B, 117
Schneider, Edgar W, 119
Schróder, Anne, 121
Schweers, C William, 183

Name index

- Schwegler, Armin, 134–136, 142, 147, 150
Schwenter, Scott A., 135, 136
Sebba, Mark, 119
Sebro, Louise, 15, 17, 18, 26, 31, 43, 49, 52
Seidel, Nathanael, 2, 15, 42
Sensbach, Jon F, 24
Sessarego, Sandro, 4, 139, 150
Sharp, William F, 138, 140
Shenk, Elaine, 183
Silva, Adriana Maria Paulo da, 158, 161
Silverstein, Louise B, 188
Simonsen, Gunvor, 14, 16–18, 24, 25, 29, 49, 54
Singler, John Victor, 15
Sippola, Eeva, 119
Slobin, Dan, 3, 65, 66, 68, 69, 91
Souza, Marina de Mello e, 161
Stahl, Ann B, 35
Stavans, Ilon, 179
Stein, Peter, 20
Stern, Nancy, 178
Stolz, Thomas, 20

Talmy, Leonard, 2, 65–68, 72, 84, 91–93
Tamargo, Rosa E. Guzzardo, 183
Tan, Daniel, 203
Tan, Peter K, 203
Thomason, S, 183, 191
Thornton, John K, 44
Tomei, Renato, 108
Torres, Lourdes, 182
Torres-González, Roamé, 181
Tupas, Ruanni, 203
Tyson, George, 18, 51

Valdman, Albert, 73, 85, 88
van der Auwera, Johan, 150
van der Voort, Hein, 19
van Rossem, Cefas, 19, 20
van Sluijs, Robert, 19, 20
Vargas Sarmiento, Patricia, 139
Vélez, Jorge A, 183
Velupillai, Viveka, 21, 22
Vigouroux, Cécile B., 99
Vos, Jelmer, 26
Voyages Database, 140

Walsh, Maureen., 180, 184
Warnke, Ingo Hans, 119
Weindl, Andrea, 20
Wells, John Christopher, 112
Werner Cantor, Erik, 139
Westergaard, Waldemar, 14, 25
Westermann, Diedrich H., 36
Williams, Caroline A, 139
Williamson, Kay, 23
Winer, Lise, 119, 122
Wolfram, Walt, 204

Yumoto, Kazuko, 179

Zabrodskaia, Anastassia, 179
Zentella, Ana Celia, 200
Ziober, Fernanda Maciel, 165

New research on circum-Caribbean creoles and language contact

This volume features research papers dealing with creolized and partially restructured language varieties in the wider Caribbean region. Initially conceived of as a conference volume drawing on papers presented at the 2017 Summer conference of the Society for Pidgin and Creole Linguistics, organized by the Universities of Tampere and Turku, Finland, the authors have since expanded the content of their original papers substantially, contributing to the empirical and analytical depth of their submissions. The volume ultimately aims both to validate new contact language research with this regional focus, as well as to stimulate further research on the fascinating language varieties that have developed and continue to thrive in the Caribbean region.