

DE GRUYTER  
MOUTON

*John H. McWhorter*

# LINGUISTIC SIMPLICITY AND COMPLEXITY

WHY DO LANGUAGES UNDRRESS?

LANGUAGE CONTACT AND BILINGUALISM

DE  
|  
G

# Linguistic Simplicity and Complexity

## Why Do Languages Undress?

*by*

John H. McWhorter

De Gruyter Mouton

ISBN 978-1-934078-37-2  
e-ISBN 978-1-934078-40-2  
ISSN 2190-698X

*Library of Congress Cataloging-in-Publication Data*

McWhorter, John H.

Linguistic simplicity and complexity : why do languages undress? / by John H. McWhorter.

p. cm. — (Language contact and bilingualism; 1)

Includes bibliographical references and index.

ISBN 978-1-934078-37-2 (alk. paper) — ISBN 978-1-934078-39-6 (pbk. : alk. paper)

1. Complexity (Linguistics) 2. Second language acquisition. 3. Languages in contact. I. Title.

P128.C664M39 2011

417'.22—dc23

2011018843

*Bibliographic information published by the Deutsche Nationalbibliothek*

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

© 2011 Walter de Gruyter GmbH & Co. KG, Boston/Berlin

Typesetting: Apex CoVantage, LLC

Printing: Hubert & Co. GmbH & Co. KG, Göttingen

☉ Printed on acid-free paper

Printed in Germany

[www.degruyter.com](http://www.degruyter.com)

## Acknowledgments

The chapters in this book other than the Introduction began as the following articles in order. Deepest thanks to the editors concerned for allowing me to reproduce them in this volume. They were:

- Tying Up Loose Ends: The Creole Prototype After All. *Diachronica* 28: 82–117 (2011)
- What the Creolist Learns from Cantonese and Kabardian (Review article of *Phonology and Morphology in Creole Languages*, ed. by Ingo Plag). *Diachronica* 23: 143–184. (2006)
- Review article of *Deconstructing Creole*, ed. by Umberto Ansaldi, Stephen Matthews and Lisa Lim. *Journal of Pidgin and Creole Languages* 23: 289–306. (2008)
- Oh, Nóó! A Bewilderingly Multifunctional Saramaccan Word Teaches Us How a Creole Language Develops Complexity. *Language Complexity as an Evolving Variable*, ed. by Geoffrey Sampson, David Gil and Peter Trudgill, 141–163. Oxford: Oxford University Press. (2009)
- Hither and Thither in Saramaccan Creole. *Studies in Language* 32: 163–195. (2008)
- Complexity Hotspot: The Copula in Saramaccan and Its Implications. *Proceedings of the Workshop on Linguistic Complexity in Interlanguage Varieties, L2 Varieties, and Contact Languages*, ed. by Bernd Kortmann and Benedikt Szmrecsanyi. Berlin: Mouton De Gruyter. (2011)
- Why Does a Language Undress?: Strange Cases of Indonesia. *Language Complexity: Typology, Contact, Change*, ed. by Matti Miestamo, Kaius Sinnemäki and Fred Karlsson, 167–90. Amsterdam: John Benjamins. (2008)
- Affixless in Austronesian: Why Flores is a Puzzle and What To Do About It. *Austronesian Undressed*, ed. by David Gil, Scott Paauw and John McWhorter. Berlin: Mouton De Gruyter. (2012)
- What Else Happened to English? A Brief for the Celtic Hypothesis. *English Language and Linguistics* 13: 163–191. (2009)

I also am grateful to Cathleen Petree and her colleagues at Mouton de Gruyter for accepting my proposal for gathering these articles into a book format, and Uri Tadmor, working at Mouton De Gruyter but also an Indonesian expert *par extraordinaire*, for helping to welcome me into the realm

of linguists exploring the contact-rich sociohistory of the Austronesian language family and supporting this volume. This book would likely not exist if it were not for a chance encounter between Cathleen, Uri and myself at a Linguistics Society of America meeting that I happened to attend. Here's to making the best of serendipity.

# Contents

Abbreviations . . . . . ix

Introduction: The creole litmus test and the *NCSL* challenge . . . . . 1

**I Creole exceptionalism**

Introduction to Section I . . . . . 19

1 The creole prototype revisited and revised. . . . . 29

2 Comparative complexity: What the creolist learns  
from Cantonese and Kabardian . . . . . 63

3 Reconstructing creole: Has “Creole Exceptionalism”  
been seriously engaged? . . . . . 103

**II Creole complexity**

Introduction to Section II. . . . . 121

4 Oh, *nóɔ!*: emergent pragmatic marking from  
a bewilderingly multifunctional word . . . . . 125

5 Hither and thither in Saramaccan Creole. . . . . 149

6 Complexity hotspot: The copula in Saramaccan . . . . . 183

**III Exceptional language change elsewhere**

Introduction to Section III . . . . . 203

7 Why does a language undress? The Riau Indonesian problem . . . . . 207

8 Affixless in Austronesian: Why Flores is  
a puzzle and what to do about it. . . . . 223

9 A brief for the Celtic Hypothesis: English in Box 5? . . . . . 261

References . . . . . 297

Index . . . . . 319

## Introduction

### The creole litmus test and the *NCSL* challenge

The chapters in this book, which with the exception of this one began as journal and anthology articles, are devoted to the exploration of a basic proposition. That is that the difference in complexity between languages' grammars is determined significantly by the extent to which second-language acquisition has played a role in their histories. This constitutes an expansion upon conceptions such as Thurston's (1987) distinction between *esoteric* and *exoteric* grammars, and Kusters' (2003: 41–4) demonstration of how amount of inflection and allomorphy varies according to whether languages are of Type 1 (mostly used as first languages) or Type 2 (mostly used as second ones).

A corollary of this framework is that grammatical simplification can be as significant a factor in languages born of second-language acquisition as grammatical mixture. That is, descriptions of creolization and other instances of language mixture as simply a matter of “feature selection” from the languages in a context are incomplete. Similarly incomplete according to my framework is a conception that any apparent simplification in a mixed language is always due to the choice of the least marked choice among the “features” available. There are cases, of which creoles are type, in which the choices made to encode most features are less “marked” than any that the source languages offer. The guiding principle, in other words, is simplification in a general sense.

#### 1. Basic assumptions

The framework these chapters argue for consists of three main planks.

1. *The normal state of language is highly complex, to an extent that seems extreme to speakers of languages like English.* I propose that the natural state of human language is the extensive marking of fine shades of semantic and syntactic distinctions, plus rampant allomorphy and irregularity – typical of languages such as those of the Caucasus and Native American languages. Languages become accreted with this degree of complexity not for functional reasons, but because normal processes of sound change, grammaticalization and reanalysis ineluctably lead to such accretion, and human cognition happens to be able to acquire such systems when the brain is young. In

## 2 Introduction

McWhorter (2007: 21–50) I outline a model of complexity that comprises the following three aspects:

*Overspecification:* Languages differ in the degree to which they overtly and obligatorily mark semantic distinctions, such as inalienable possession, the “fourth person” obviative, evidentiality, and definiteness.

*Structural elaboration:* An aspect of one grammar may differ from that aspect in another’s in terms of the number of rules (in phonology and syntax) required to generate grammatical forms; examples include morphophonemic processes, concord, and heterogenous word order.

*Irregularity:* Grammars differ in the degree to which they exhibit irregularity and suppletion.

The Nakh-Daghestanian language Lak, then, is a “normal” language in its degree of overspecification and structural elaboration:

- (1) *Insan-tura-l arcu                      žahil-minna-l                      darcun-ni.*  
men-PL-ERG money.ABS young-NMLZ.CL1.PL-ERG steal-PAST.3  
‘Young men stole the money.’

(Kazenin 2009: 401)

Lak marks ergativity; the modifier of an NP need not appear next to it, but if it does not, it takes a nominalization marker which varies according to noun class and number; the past marker varies for person. This degree of complexity is, under my analysis, that of a language that has undergone little or no second-language acquisition.

*2. Languages significantly low in this kind of complexity compared to their sisters owe this state to second-language acquisition in the past.* I propose that under ordinary conditions, languages lose a modest amount of complexity but maintain a basic level of it through the development of new complexities (cf. Chapter Seven, Section 2.2.). Under this analysis, languages that have not maintained this level have *always* been interrupted in their normal accumulation of complexity.

These languages include familiar ones such as English, the Romance languages, Persian, Mandarin Chinese, and Indonesian (McWhorter 2007). In these cases, grammatical mixture with other grammars is modest; what distinguishes them from their sisters significantly is degree of complexity. For example, in this light the “normal” Indo-European languages are Slavic, Baltic, Greek, Latin, Sanskrit, Pashto and Icelandic, while English, Dutch, the Mainland Scandinavian languages, the Romance languages and the modern Indo-Aryan languages show evidence of widespread second-language learning in the past.

I term languages of this kind Non-Hybrid Conventionalized Second-Language Varieties (NCSLs). Chapter Eight will suggest that some unusually analytic Austronesian languages of Flores and Timor also fit into this class.

3. *The languages evidencing the least complexity of all of the world's languages are creoles.* That is, where complexity has been lost to a radical degree, the language was born in a situation in which adult acquisition was universal or nearly so – namely, the languages termed creole languages. Because many creoles develop in close contact with older languages, while others are created by speakers of closely related older languages and thus maintain many complex features processible by all creators, not all languages termed creoles display the maximum degree of simplification that others do. However, there is a maximal degree of simplification that some languages display in comparison to their source languages that is unique, and all languages exhibiting this degree of simplification, despite the variation in that degree, are all ones born in rapid and widespread adult acquisition. That is, they are creoles.

As I will demonstrate via Saramaccan Creole in Chapters Four, Five and Six, creole languages of course have a degree of complexity: demonstrations of this, common since I first presented this hypothesis, do not qualify as refutations of my proposition. The point is that creoles are considerably *less* complex (according to the metric presented above) than older languages, including NCSLs, as will be demonstrated in Chapter Two.

An example is the contrast between a sentence in French and one in Haitian Creole:

(2)(a) French:

*Ils n'ont pas de ressources qui puissent*  
 3P NEG.have NEG PART resource.PL REL can.SUJ.3P  
*leur permettre de résister.*  
 3P.OBJ allow to resist

(b) Haitian:

*Yo pa gen resous ki pou pèmèt yo reziste.*  
 3P NEG have resource REL can allow 3P resist

(Ludwig, Telchid, & Bruneau-Ludwig 2001: 164)

The framework in this book cannot accommodate a claim that French's inflectional morphology, including irregularities such as *ont* and *puissent*, plus case-distinguished pronominal forms such as *ils* vs. *leur*, the partitive marker *de* and its very particular conditioning, and the occurrence of object pronominal *leur* in the clitic-climbing position before the verb, does not entail

that more rules need to be applied to generate the French sentence than the Haitian one in which all of these traits are absent. Note also: I do not designate the plural *-s* on French's *ressources* as indicating more overspecification than Haitian's number-neutral *resous*, as it is not pronounced. Yet, morphophonemic liaison processes often preserve plural *-s* in spoken French whereas this process is absent in Haitian.

Equally unsuccessful would be an investigation as to whether Haitian exhibits complexity elsewhere in its grammar that leads it to match French in degree of complexity (cf. McWhorter 2005a: 102–41 on the comparison of Fongbe with Saramaccan and Haitian, and McWhorter 2007: 36–50 on such a comparison of Estonian and Saramaccan). The overall grammar of Haitian is markedly less complex than that of French, and the framework of these chapters will treat this as symptomatic of unusually prevalent adult acquisition.

## 2. Why these assumptions require statement

Few linguists would disagree with the basic proposition that second-language acquisition entails grammatical simplification.

However, there is a resistance in some quarters to the idea that such simplification could leave its mark on languages over time other than creoles. The languages I term NCSLs are commonly treated as having fallen into their lighter degree of complexity as a matter of chance (cf. Thomason & Kaufman 1988: 263–342 on English), or are attributed to a “drift” towards analyticity in Europe and the Middle East (the common take on Romance vs. Latin, or the colloquial Arabics vs. Classical). This leads to questions. What was it about earlier ages that conditioned the opposite “drift” into the complexity of ancient languages like Sanskrit, Latin, and Classical Arabic? What has made all of the subsequent stages of these languages “drift” towards analyticity together? And why no “drift” towards analyticity in Native American, Australian Aboriginal, Bantu, Uralic, Palaeosiberian, Dravidian or, actually, most languages on earth?

I propose that widespread second-language acquisition, an accepted source of grammatical simplification, was the cause. To wit, second-language acquisition can have effects on grammars more permanent than mere matters of individual speakers' interlanguage and post-critical stage fossilization, and less extreme than pidginization or creolization.

Meanwhile, while the influence of second-language acquisition on creoles is readily apparent to creolists, there is a resistance to the idea that

it left a decisive enough impact on them to distinguish them from other natural language grammars. Plag (2008a, 2008b) argues carefully that creoles are indeed “conventionalized interlanguages of an early stage,” but nevertheless rejects the entailment – which would seem inherent – that creole language grammars are simpler than older language grammars (cf. Plag 2008a: 130–1). This demonstrates an acute discomfort, typical among creole specialists, with the characterization of creoles as “simpler” (cf. Kouwenberg 2010: 173: “McWhorter’s arguments that prototypical creoles converge on a set of properties which he designates ‘simple’ have won him few friends in the field”). The consensus would appear to be that whatever the effects of second-language acquisition on creoles, older languages can also fall into the same lesser degree of overall complexity as a matter of chance.

However, the languages in question have yet to be identified. There have indeed been a few attempts to specify older languages that display the three features of the Creole Prototype outlined in McWhorter (2005a: 9–37), as discussed in Chapter One. None of these go through, however, for reasons outlined in that chapter. In any case, the Creole Prototype is not intended as a condition of maximum simplicity, despite frequent misunderstandings on that score as evidenced, for example, by Kouwenberg’s statement above. There is nothing inherently complex about tone or noncompositional derivation, for example. The Prototype refers to these features because of their being symptomatic not of complexity, but of a language’s age.

The disproof that the least grammatically complex languages are creoles would be a language with no history of widespread adult acquisition that displayed as little complexity (according to the metric above) as a subset of creoles do. It is reasonable to hypothesize that such a language does not exist.

### 3. The Creole Litmus Test

In this chapter I will apply this framework to a little-known language, in order to illustrate the basic assumptions which will be examined more closely in the subsequent chapters.

I claim that creole languages are synchronically identifiable, even without recourse to information about their histories. That is, a creole is “a kind of language,” not just a sociohistorical term. My stipulation – simplified here; cf. Chapter One for a more detailed presentation – is that a subset of languages show the following three kinds of indication that they were born recently of pidgins:

*phonological*: little or no use of tone to distinguish monosyllables or grammatical categories

*morphosyntactic*: little or no inflectional morphology

*semantic*: little or no noncompositional combinations of derivational markers and roots.

This proposition has been highly controversial among creole specialists, and in the years since I first proposed it, a conclusion would appear to have set in that, as Singler (2006: 159) puts it, “No linguistic litmus test of creole status exists – or could exist.” However, I believe that one indeed could and does. The death knell sounded by Singler, as well as Ansaldo, Matthews & Lim (2007) and others is premature.

Chapter One will present a revision of the Creole Prototype hypothesis which will make its theoretical motivation clearer, and show that certain data thought to constitute exceptions to it in fact do not. Here, however, I will give a preview of how the Creole Prototype indeed constitutes a litmus test for creolization, with an examination of an isolating language that has not been considered a creole under any definition. This is a kind of language that could be seen as suggesting that even older languages can wend into a creole-like typology, and that there is therefore properly no synchronic essence limited only to creole languages. However, the language in question, isolating though it is, gives ample signs of its antiquity.

New Guinea, home to hundreds of languages spoken by small groups, is a setting that typically yields “normal” grammars in the sense specified above: robustly complex, most saliently in considerable morphological elaboration on the verb (Foley 1986: 12). An example is Amele:

- (3)(a) *Ija hu-m-ig sab j-ig-a.*  
 I come-SP-1S food eat-1S-RPAST  
 ‘I came and ate the food.’
- (b) *Ija ho-co-min sab ja-g-a.*  
 I come-DFP-1S food eat-2S-RPAST  
 ‘I came and you ate the food.’

(Roberts 1988)

Amele displays various typical features of the languages termed, despite the multiple families it comprises which are unlinked to a reconstructed ancestor, “Papuan.” SP and DFP stand for “same pivot” and “different pivot,” the latter of which is a switch-reference marker. The subject is indexed with an affix on the verb corresponding to its person and number, which in addition differs according to the pivot marker. The *come* root is variable according to

its affixation; such roots often vary even suppletively in Papuan languages. The past marker in these sentences refers only to the day in question, as opposed to a more distant past.

Things are quite different in Abun, of West Papua in the Northern Bird's Head peninsula. Abun is a rare thing: a largely isolating Papuan language:

- (4) *Men ben suk no nggwe yo, men ben suk sino.*  
 we do thing LOC garden then we do thing together  
 'If we do things at the garden, then we do them together.'

(Berry & Berry 1999: 23)

Isolating structure hardly bars extreme complexity, in itself. To assess whether Abun is identifiably less grammatically complex than a language like Amele, we must check whether it has free morpheme equivalents to the switch-reference markers or allomorphically variant subject-indexed suffixes. It, in fact, does not. Plus, not only is there no particle dedicated to the specific gradation of "earlier today" as there is in Amele – Abun does not mark tense overtly at all!

Here, then, is a language as isolating as a creole like Haitian – and, unlike isolating languages like Chinese and Yoruba, we see in (4) no tonal markings. In fact, Abun does have three phonemic tones. However, their functional load is quite light and Abun's grammarians venture that the distinction is on its way to extinction. Cases of three-way contrasts on monosyllables are rare: Berry & Berry (1999: 21) present but one, *šè* "flow," *šé* "flood," *še* "big." Even two-way contrasts are rare (*ndò* "good," *ndó* "bitter"), uncommon beyond an only slightly productive usage to indicate plurality (*an* "he, she, it," *án* "they," *ndam* "bird," *ndám* "birds" [ibid. 21]).

Yet, it would be infelicitous to treat this modest degree of contrastive tone as disqualifying Abun as a creole, because some creoles have marginally contrastive usage of tone as well. Saramaccan has *dá* "give" vs. *da* "be," *á* "not" vs. *a* "he, she, it," the future marker *ó* vs. the affective marker *o*. Principense Creole Portuguese has an HH pattern for some deverbal nouns contrasting with an LL one for the verbs, which does not correspond to the stress patterns: (*fálá* "speech," *fàlà* "to speak," *témá* "insistence," *tèrà* "to insist" [Maurer 2009: 26]).

In truth, all of these cases can be traced conclusively to developments after the genesis of the creoles, as will be discussed in Chapter One (which will also address similar data from Papiamentu). They are signs of the gradual accretion of complexity diagnostic of a language's transformation over time, but were not products of initial creolization itself. However, to the extent that this will seem a matter of special pleading to some – although it isn't! – we will proceed as if Abun's degree of tone is immaterial to assessing whether it is a creole.

In the same fashion, the fact that Abun has a transitivizing suffix on verbs (although it is used only on some) cannot disqualify it as a creole, given that Tok Pisin and its sister languages are readily admitted to the class despite departing from the Prototype with their transitivizing *-im* suffix (*Mi boil-im wata* “I boil water”):

- (5)(a) *Ji nyu.*  
I fear  
‘I am afraid.’
- (b) *Nu nyu-wa men o nde.*  
2P fear-TR 1P again NEG  
‘Don’t fear us anymore.’

(Berry & Berry 1999: 27)

The distinction of ancient languages from creoles need not, then, rely on hair-splitting. Despite “flutter” such as the minimal tone and the transitive marker, there are two aspects of this isolating language that reveal it quite readily as ancient rather than as born of a pidgin recently – i.e. that disqualify it as a creole.

One is that as we would expect of an ancient language, combinations of derivational morphemes and roots have often drifted into noncompositionality. The noncompositionality in question is not the typical slight drifts from strict compositionality that derivation-root combinations undergo in all languages as the result of cultural conventions, such as the use of *transmission* to refer to an automobile part. Rather, what is diagnostic of antiquity is such combinations whose meanings are much more idiosyncratic and inexplicable synchronically than this, outright lexicalizations such as English’s *understand*, or its verb-particle combinations such as *make out* or *put up with*.

In Abun, this kind of noncompositionality occurs in fact with the transitive suffix *-wa*: *kon* “cook,” *kon-wa* “celebrate;” *bi* “give,” *bi-wa* “pay for;” *ki* “say,” *ki-wa* “ask for” (ibid. 29). Or, the original meaning of *ket* is directional, roughly “over to”:

- (6) *Men tot ket nden.*  
1P cut LOC path  
‘We cut across to the interior path.’

However, combined derivationally with roots, it can yield highly unpredictable meanings: *ye* “difficult,” *ye-ket* “surprised.”

The nature of such derivational morphemes in Abun can be demonstrated by their combination with the one verb *ki* “to say,” yielding a wide range of idiosyncratic meanings:

<i>wa</i> “to, for”	<i>ki-wa</i> “ask for”
<i>ket</i> “to”	<i>ki-ket</i> “slander”
<i>bot</i> “through”	<i>ki-bot</i> “discuss”
<i>gat</i> “join”	<i>ki-gat</i> “speak persuasively”

The second Abun trait that reveals it as antique is that while it is isolating, it is nevertheless inflected. That is, it does not have bound inflection, but it has free morphemes that qualify as Inflection in a larger grammatical sense, along the lines specified by Kihm (2003). Creoles, too, have free morpheme Inflection: their tense-marking “particles,” for example, are inflections which differ from inflectional affixes in the same function only in how they are handled by syntax (they Move rather than Merge, in Kihm’s framework).

However, one aspect of the revision of the Prototype formulation in Chapter One is that languages born recently of rudimentary pidgins do not have, either bound or unbound, *paradigmatic* inflection: that is, allomorphic batteries of noun class markers, conjugation class markers, etc. Paradigmatic morphology is absent in a rudimentary pidgin, and from there, can only develop via grammaticalization or reanalysis over time. Abun, however, has paradigmatic morphology in the form of its battery of numeral classifiers. Such classifiers are analogous to noun class or gender markers (cf. Grinevald & Seifart 2004):

<i>bo</i>	fruit, motor
<i>but</i>	bundles
<i>ge</i>	person, animal
<i>gwes</i>	bamboo (cut pieces)
<i>is</i>	tuber
<i>ka</i>	person
<i>ke</i>	tree, house
<i>koi</i>	stick (cut pieces)
<i>sak</i>	cloth
<i>wak</i>	cloth

(Berry & Berry 1999: 42)

Therefore, the way that Abun reveals itself as ancient despite having almost no affixation and little contrastive tone is in having 1) noncompositional derivation, and also 2) “Inflection” of a kind – paradigmatic – that is absent in creoles. This absence makes sense if creoles were born a few centuries ago as makeshift pidgins created by adults. This is a reason to suppose that creoles were indeed born in this way – on top of the sheer empirical fact that relatively new creoles which emerged amidst ample documentation can be seen over time unequivocally emerging from rudimentary pidgins; namely, Tok Pisin and its sister creoles of Oceania.

Unsurprisingly, Abun also departs from creoles in having some over-specifications of a kind rarely encountered in them (cf. McWhorter 2005a: 72–101), such as a distinction between alienable and inalienable possession (Berry & Berry 1999: 79).

The numeral classifier issue is important. In earlier work I have stressed that noncompositional derivation crucially distinguishes analytic and toneless older languages from creoles (cf. McWhorter 2005a: 16–18). However, numeral classifiers do as well, as they are typical in Southeast Asia which contains some of the only languages in the world with neither inflectional affixes nor tone, and are perhaps a handier metric. Few grammars devote much space to noncompositional derivation given that it is not, properly, grammar. Moreover, the alphabetical order of dictionaries means that they can only be searched easily for derivational morphemes' semantic contributions when the morphemes are prefixes, rather than suffixes. Numeral classifiers, like inflectional affixes and tones, are readily described even in briefer treatments.

Thus there can be, at present, no conclusive claim that a litmus test for creole status does not exist or would be theoretically impossible. Parkvall (2008) indicates that creoles are measurably less complex than older languages according to grammatical features tabulated in the *World Atlas of Language Structures*. Especially useful in this light is Bakker, Daval-Markussen, Parkvall and Plag's (2011) two comparisons. One is of creole languages and a range of older languages with "creole"-like profiles such as analyticity, according to the presence or absence of a substantial list of grammatical features examined in, again, the *World Atlas of Language Structures*. The other is of creoles and older languages according to the features tabulated in a wide range of creoles by Holm & Patrick's (2007) project comparing a range of creoles. Bakker, Daval-Markussen and Parkvall find that when the data is submitted to SplitsTree software, creole languages group together as a class – despite the sample including several analytic languages and African languages, two categories which we would expect to cluster with creoles according to the idea that there exists no synchronic class of creole language. All statements that there exists no synchronic distinction between creoles and older languages is, as of the publication of these articles, invalid without engagement with them.

Importantly, one kink in Parkvall and Bakker's result is that Hmong, a highly isolating Sino-Tibetan language, groups with the creoles. However, this is because they happen not to include numeral classifiers in their tabulation, these being rarely discussed as relevant to creoles because of their absence in them (as well as, probably, in the West African substrate languages associated with them). However, if classifiers were included, then Hmong, which has over 50, would no longer cluster with the creoles.

#### 4. Abun as a NCSL

Even if Abun is identifiable as not the product of the recent pidginization and reconstitution of a language, the fact remains that it is notably less grammatically complex than the typical Papuan language. This is true not only in its analyticity but in overall degree of grammatical elaboration of the kind specified above, in a fashion that does not lend itself to quantification, but is starkly apparent nonetheless. Berry & Berry's (1999) Abun grammar reveals a language requiring a much briefer description than languages like Amele, Nasioi, Yimas, Iatmul, and the vast majority of the hundreds of Papuan languages – i.e. of the kind that might lead to a preliminary impression that it “seems like a creole” or is “creole-like.”

##### 4.1. The nature of the question

The question as to why is perhaps uninteresting from the perspective of the synchronically oriented linguist. To him or her, some languages are analytic, some are not, but all operate according to grammatical principles. From a diachronic perspective, however, the question is urgent.

Languages like Abun are sometimes approached according to an assumption that they simply “lost their morphology” as if this were a documented regular process of language change. However, it is not.

Why, after all, would a grammar over time lose not just some, but virtually all of its morphology? That is, precisely what would the process be? That such a thing seems even initially plausible to many linguists is perhaps an artifact of how English and some other Western European languages happened to develop. However, first, English and the Romance languages retain a goodly amount of inflectional and derivational morphology nonetheless, remaining fusional languages. Second, as noted above, even these developments are atypical of language change and constitute a challenge to diachronic theory (as addressed in McWhorter 2007). For example, under an assumption that losing almost of its morphology is an unsurprising pathway for a grammar to take, the languages of Australia are a stunningly intractable puzzle – there is not a single language on the continent that has somehow shed its affixation.

Moreover, we cannot propose that Abun never had affixation. Affixation develops as the result of ordinary processes of grammaticalization and re-analysis which are universally operant in human language. It is inconceivable that a human language could exist for as long as the 65,000 years often reconstructed as the length of human occupation of New Guinea without developing affixes of some kind. In fact, after this much time, we see from