Thomas B. Klein

Creole phonology typology: Phoneme inventory size, vowel quality distinctions and stop consonant series

1. Introduction¹

Most practitioners in the field of linguistics would agree that the phonology of Creole languages is considerably understudied (see for example Singh and Muysken 1995), notwithstanding the favorable attention it has received in recent years (see in particular the contributions in Plag 2003). Most of the existing work centers on a small number of themes. One of the core concerns is the comparison of Creoles with their European lexifiers and their non-European substrates. Another focus is on the phonological development from pidgin to Creole stages or within the history of a given Creole or a small group of Creoles. Synchronic work on Creole phonology tends to be couched in variationist frameworks or is part of structural descriptions where it is often given fairly short shrift. The geographical focus of surveys discussing the structures of Creole languages tends to be on either the Atlantic (see, e.g., Holm 2000, Parkvall 2000) or the Pacific Creoles (see, e.g., Mühlhäusler 1997). Furthermore, only Creoles with certain Indo-European lexifiers are usually considered.

Generalizations concerning the phonology of Creole languages are invoked quite frequently, but they tend to be based on a fairly narrow sample of languages and are not grounded in a typological analysis. For instance, Bender (1987) suggests a list of probable phonological universals based on "overall impression", but acknowledges that a comparative study of Creole phonologies is "obviously one of the most pressing needs in Creole studies" (Bender 1987: 42). The present study is designed in part to address this need.

It has been claimed that the linguistic structure of Creole languages is more alike or simpler than that of non-Creole languages. Influential concepts such as Bickerton's bioprogram (see Bickerton 1981 *et seq.*) are based on the seeming surface similarity of Creole languages.² Substratist and superstratist thinking relies to a good degree on the idea that a given set of structures is more or less iterated across Creole languages, thus making them

¹ Earlier versions of parts of this material have been presented at the University of Edinburgh, at the University of Essex, at the 2003 SPCL meeting in Atlanta and at the Second International Workshop on the Phonology and Morphology of Creole Languages in Siegen. Thanks to the audiences there and in particular to Jacques Arends, Parth Bhatt, Robert Clark IV, Stuart Davis, Shelome Gooden, Mya A. Green, Meta Y. Harris, Winford James, Bob Ladd, Miriam Meyerhoff, Mits Ota, Peter Patrick, Ingo Plag, Norval Smith and two anonymous reviewers for useful comments, discussion and encouragement. This research has been supported in part through a Faculty Research Grant at Georgia Southern University in the summer of 2003, which is hereby gratefully acknowledged. All responsibility for errors lies with the author.

² Note, however, that the evidence in support of the bioprogram hypothesis was drawn from morphosyntax, not phonology.

appear more similar than the linguistic diversity in the formative contact situation would suggest. The idea that Creole languages are grammatically simplified in relation to their source languages, especially the superstrates, is considered a truism by many, professional linguists and laypersons alike. Others have argued that the linguistic structure of Creole languages is simple in some absolute sense, not just as compared to the languages in the contact situation, but also in comparison to non-Creole languages in general.

This idea of absolute simplicity has been advocated prominently in recent work. In particular, it has been claimed that the world's simplest grammars are Creole grammars (McWhorter 2001a, b), henceforth referred to as the *Creole simplicity hypothesis*. According to this hypothesis, Creoles display significantly less complexity as a group than the rest of the world's natural grammars. This entails that there would be no significant set of non-Creole languages with simpler grammars than Creole languages. McWhorter (2001a) invokes markedness in the sense of a Greenbergian implicational metric and inventory size in his discussion of phonology. To a significant degree, the phonemic and tonal inventories of Creole languages would employ fewer distinctions and would contain less marked elements by this measure.

Markedness and inventory size are to a good degree interdependent, but they are also distinct parameters. The implicational relationship between unmarked and marked segments affects the quantity dimension because marked segments usually occur in relatively large inventories that already possess the unmarked segments. However, this is not necessarily the case. For instance, Rotokas, the language with the smallest inventory of eleven phonemes in Maddieson's (1984) typological analysis, includes two highly marked segments, the alveolar tap and the voiced bilabial fricative, as part of its inventory of six consonantal phonemes (see Maddieson 1984: 367). In other words, the Rotokas case demonstrates the potential distinctness of markedness and inventory size because its very small set of phonemic consonants contains marked segments nonetheless.

Meaningful discussions and interpretations of the sound structure of Creole languages are hampered by the lack of comprehensive approaches to segmental and suprasegmental markedness and the absence of wide-ranging typological analyses of Creole phonology. Detailed treatment of markedness is beyond the scope of the present work. Instead, the focus of this paper is on the size of phonemic inventories, the number of vowel quality contrasts and the number of contrastive stop series. A quantitative typological investigation is undertaken to illuminate these parameters in a systematic sample of Creole languages. The results from this investigation are compared to the typology of non-Creole languages as laid out in Maddieson (1984). No comparison of Creoles to their lexifier languages is made. Instead, Creoles are compared to the entire typological sample of non-Creole languages. Creole sound systems are not primarily compared to non-Creole sound systems with regard to what kinds of phonemes they include but how many phonemes Creoles employ in comparison to non-Creoles along a given parameter.

The present analysis focuses on Creole phoneme inventories and the question if the phonological inventories of Creoles are quantitatively simpler or more alike than those of non-Creoles. The aims of the paper are to construct and analyze a typological sample of phoneme inventories from the range of Creoles around the world and to launch a systematic quantitative study of the phonemes of Creole languages with those of non-Creole languages. Maddieson's (1984) ternary scale, *simple – typical – complex*, to measure the phonological complexity of non-Creole languages is of particular importance in this regard.

Creole phonology typology

Based on this scale, the results of this paper support the idea that Creole sound systems are quantitatively very typical sound systems, referred to henceforth as the *Creole typicality hypothesis*. According to this hypothesis, Creole languages exhibit typical phonological systems that center on typological middle ground.

The remainder of this paper is structured as follows. I show next how the present typological sample is constructed. The analyses of phoneme inventory size, vowel quality distinctions and stop phoneme series are presented in turn. The implications of the typological findings are discussed in each section with particular reference to the question of simplicity versus middle ground in Creole phonology before some conclusions are offered.

2. A typological sample of Creole languages

Maddieson's (1984) erudite and influential work on the typology of phonemic systems is the empirical and methodological backdrop for the current study. Maddieson's research is based on the UCLA Phonological Segment Inventory Database (UPSID). This database contains 317 languages, none of which is a Creole. Maddieson employs a quota sample to enable sampling efficiency and to ensure maximal genetic diversity in the database. He allows no more than one language from any small family grouping; for instance, West Germanic and North Germanic languages are represented through one language each.

The present sample is constructed to represent the range of Creole languages around the world. Care is taken to balance Creoles from diverse locations and with European and non-European lexifiers. If Maddieson's quota of one language per small sub-grouping were to be applied, the number of Creole languages in the resulting sample would be too small for the database to be truly meaningful. For instance, only one French-lexified Creole from the Atlantic area could have been included according to this logic. Thus, to ensure balance, diversity and sampling effectiveness, two Creoles with the same lexifier are considered from a given geographical area wherever possible.

Acrolectal Creoles or varieties are generally excluded. The reason is that they are closest to and, hence, least distinct from the corresponding lexifier languages. The focus is on mesolectal and, in particular, basilectal varieties. One consequence regarding Atlantic English-lexified Creoles is the inclusion of Suriname Creoles at the expense of Caribbean island Creoles, given that the latter tend to be closer to the lexifier. For analogous reasons, Gulf of Guinea Portuguese-lexified Creoles have been included, whereas Cape Verdean Creole varieties are not considered. This move is not meant to impart judgment on the ultimate degree of 'Creoleness' of these languages, however. Recent additions to the phoneme inventory of a given Creole through loans from European lexifiers have been avoided wherever they are obvious or the sources identify them. For example, front rounded vowels from Dutch loans in Papiamentu have been disregarded to focus on the core rather than the periphery of the phoneme system. There is also a bias towards the quality of the available descriptions. Creole languages are chosen in part based on how detailed the information on the phoneme system is in the sources and how reliable this information seems. The current sample contains twenty-three (23) Creole languages. Table 1 lays out how they are classified according to geographic area and lexifier language and which sources have been used to gather the phonological information. Note that the sample includes varieties that have only recently become nativized as Creoles such as Tok Pisin and Sango.

Area	IE-Lexifier	Creole	Sources
Atlantic	Dutch	Berbice	Kouwenberg 1994
		Negerhollands	Sabino 1990, Stolz 1986
	English	Ndyuka	Huttar and Huttar 1994
		Saramaccan	Aceto 1996, Bakker et al. 1995, Rountree 1972
	French	Haitian	Corne 1999, Lefebvre 1998, Tinelli 1981
		St Lucian	Carrington 1984
60	Portuguese	Angolar	Lorenzino 1998, Maurer 1995
		São Tomense	Ivens Ferraz 1979
	Spanish	Papiamentu	Kouwenberg and Murray 1994, Kouwenberg and Muysken 1995, Maurer 1998, Munteanu 1996
		Palenquero	Bickerton and Escalante 1970, Lewis 1970, Megenney 1986, Patiño Rosselli 1983, 1999
Indian Ocean/	German	Unserdeutsch	Volker 1982
Pacific	English	Bislama	Meyerhoff 2003, Tryon 1987
		Tok Pisin	Laycock 1984, Mühlhäusler 1984, Smith 2002, Verhaar 1995
	French	Mauritian	Baker 1972
	1	Тауо	Corne 1999, Ehrhart 1993
	Portuguese	Kristang	Baxter 1988
		Sri Lankan	Smith 1977
	Spanish	Zamboangueño	Forman 1972, Whinnom 1956

Table 1: Creoles in sample

Area	Non-IE-Lexifier	Creole	Sources
Africa	Bantu	Kituba	Fehderau 1962, 1966, Institut National 1982, Mufwene 1997, Nida 1956, Swift and Zola 1963
		Sango	Pasch 1997, Walker and Samarin 1997
	Arabic	Juba Arabic	Kaye and Tosco 2001, Owens 1997
		Nubi	Pasch and Thelwall 1987, Owens 1997
Asia	Malay	Baba Malay	Lim 1981, Pakir 1986

Table 1 cont.:

The lexifier languages are assigned the central role in classifying the Creoles in the sample linguistically. The basic geographical division for Creoles with Indo-European lexifiers is between Atlantic creoles and the parts of the world delineated by the Indian Ocean and the Pacific. The five European lexifier languages Dutch, English, French, Portuguese and Spanish are easily matched with two Creoles each in the Atlantic area. This is less straightforward for the Pacific Rim/Indian Ocean Creoles. There is no Dutch-lexified Creole in this area, so Rabaul Creole German (Unserdeutsch) takes its place given that the lexifier languages are very closely related. As far as I am aware, there is no source comparable to Forman's (1972) dissertation on Zamboangueño that would discuss a second Pacific Spanish-lexified Creole in enough detail to extract the necessary phonological information. Creoles with non-Indo-European lexifiers are considered equally important for the comprehensive understanding of Creole languages. They are classified geographically according to the continent on which they are found. There is only one Malay-lexified Creole represented because of a lack of descriptive resources to extract reliable phonological information on another language of this type.

3. The size of Creole phoneme inventories

This section serves to compare the results of Maddieson's (1984) investigation of phoneme inventory sizes with the current database of Creoles. Maddieson's practice of representing each segment considered as phonemic by its most characteristic allophone has been followed throughout the present study. One problematic area in determining the segment inventories involves choosing between a unit or a sequence interpretation of elements such as affricates, prenasalized stops, and diphthongs. Maddieson examined the available evidence "with some prejudice in favor of treating complex phonetic events as sequences (i.e. as combinations of more elementary units)" (p. 6).

Maddieson's (1984) survey has uncovered a great range in the size of the phoneme inventories of non-Creole languages. The smallest inventories of 11 phonemes are found in

Rotokas and Mura, whereas the largest inventory of 141 phonemes is found in the Khoisan language !Xũ. However, the typical phoneme inventory size is between 20 and 37 segments. According to Maddieson, 70% of the UPSID languages, that is 222 out of 317 languages, fall into this category. Inventories are classed as simple by Maddieson when they are smaller than 20 phonemes, whereas inventories with more than 37 phonemes are classed as complex. Thus, this metric employs the tripartite distinction of simple, typical, and complex. This measure of complexity is applied analogously to the Creoles to enable a straightforward comparison with non-Creole languages.

When variations of the size of the phoneme inventory of a given Creole are found in the descriptions, certain general guidelines have been applied. Whenever there appear to exist marginal phonemes with limited distribution in a Creole language, these are excluded from the present database. The idea is to represent the core of a given inventory and to abstract away from its periphery. The issue of the interpretation of consonants as units versus sequences has been treated as follows. Affricates, labiovelar stops and homorganic prenasalized stops are always treated as units; non-homorganic nasals plus obstruents are treated as sequences and, hence, are not part of the inventory count unless the descriptive literature on the Creole in question makes a case for the treatment as units. Diphthongs are treated as sequences, given that diphthongs in Creole languages generally exhibit vocoid qualities that are also found in the monophthongs. For example, a diphthong aj is treated as a sequence of the vowel phonemes a and i. Nasal vowels are counted as distinct from oral vowel phonemes in determining the inventory size. If the status of certain phonemes such as nasals is doubtful given the information in the descriptive literature, they are not included in the inventory count. The overall idea is to employ a conservative count of the size of Creole inventories. Nonetheless, the emphasis is on what Creoles have, not what they may lack in comparison to other types of languages.

The number of the phonemes in the Creoles surveyed is displayed in Table 2 below. The table is evidence that Creoles display a narrow range in the size of their phonemic inventories. The smallest inventory has 19 phonemes, whereas the largest one has 37. Recall that the range for non-Creole languages is much wider, from 11 to 141. The vast majority of Creole languages exhibits the typical non-Creole inventory size of 20 to 37 phonemes, following Maddieson's measure. Only one simple inventory is found; 96% (22 of 23) of Creoles show the typical size, whereas no Creole inventories are complex by Maddieson's measure. Given these figures, the frequency of typical inventory sizes in Creole versus non-Creole languages may be calculated statistically. Given that the numbers at issue are quite small, the Fisher's exact probability test (Langsrud 2004, Lowry 2000) or the chi-square test with Yates' correction for continuity (Preacher 2003) may be used to calculate statistical (non-) significance. The Fisher's exact test shows that the frequency of Creole languages with the typical inventory size is significantly higher than in the UPSID languages (p < 0.007; two-tailed). This result is duplicated using Yates' chi-square (5.86; p < 0.02). This means that significantly more Creole languages have phoneme inventories of typical size than non-Creole languages. In other words, Creole phoneme inventories are quantitatively more middle-of-the-road than the UPSID inventories.

Creole	No. of phonemes
Ndyuka	19
Kituba	20
Tok Pisin	21
Berbice, Juba Arabic, Tayo	22
Palenquero	23
Bislama	24
Baba Malay	25
Mauritian, Nubi, Sri Lankan	26
Negerhollands, Unserdeutsch, Kristang, Zamboangueño	27
Papiamentu	30
São Tomense	31
Haitian, Saramaccan ³	32
St Lucian	33
Sango	35
Angolar	37

Table 2: Creole phoneme inventory size

The inventory sizes of Creoles are also very evenly distributed within their range. The mean number of segments in a Creole phoneme inventory is just under 27, whereas the median is 26 segments. The small differential in mean versus median points up the even distribution. Maddieson reports that the mean number of phonemes per language in the UPSID database is a little over 31; the median falls between 28 and 29. These numbers are an outcome of the greater range in diversity in the size of non-Creole inventories. The model number of consonants in an UPSID inventory is 21 (Maddieson 1984: 12), whereas a system of 5 vowel qualities is most frequent (see also below). Interestingly, the mean and the median in the Creole languages correspond very closely to the sum of the most frequent number of vowel qualities (i.e., 5) and the number of the most frequently occurring individual consonant segment types in the UPSID database (i.e., 21).

The numerical results presented in this section may be discussed in light of the predictions of the competing hypotheses concerning simplicity versus middle ground in Creole grammars. Notwithstanding potential interaction with markedness, the Creole simplicity hypothesis predicts a substantial number of Creoles with small inventories of fewer than 20 phonemes. The present investigation shows that such simple inventories are very scarce. The rarity of such inventories is unexpected and surprising under the simplicity hypothesis.

Given the present measure, it emerges that Creoles do not have the simplest inventories, but the most typical ones. Nearly all Creoles fall within the range typical of non-Creole

³ According to McWhorter, there are twenty-five (25) phonemes in the inventory of Saramaccan (2001a: 139). However, descriptive sources on Saramaccan present evidence for thirty-two (32) phonemes instead (Aceto 1996, Bakker *et al.* 1995, Rountree 1972). It seems that the seven vowel phonemes of Saramaccan have been omitted from McWhorter's count.

languages clustering around an inventory size of 27. This distribution is expected under the Creole typicality hypothesis. The fact that no Creoles show inventory sizes that are complex under the current measure supports the idea that Creole sound systems may be the world's most typical sound systems. The predictions of the typicality hypothesis are also met by the findings of the investigation of the number of distinctive vowel qualities.

4. Distinctive vowel qualities in Creole languages

This section compares the results of Maddieson's (1984) study of the number of vowel qualities in phoneme inventories with the corresponding results from the Creole languages and interprets them in light of the competing hypotheses of Creole structure.

Many languages have more than one series of vowels such as long and short vowels or oral and nasal vowels. The vowels in one series can often be matched in quality with another series, so that the number of vowel phonemes is greater than the number of different vowels qualities in the languages concerned. In other languages there may be qualities in one series which do not occur outside that series, so that the total number of vowel qualities may be larger than the number found in a given series (see Maddieson 1984: 127).

This issue arises in Creole languages typically for oral versus nasal vowels. Consider Creole Sango as an instance of oral and nasal vowels matching in quality. It seems clear that Sango has the seven oral vowel phonemes / i, u, e, ε , o, o, a/ and the five nasal vowel phonemes / i, ũ, ẽ, õ, ã/ (see Pasch 1997, Walker and Samarin 1997). Given that the qualities of all nasal vowels are found in the oral vowels, Sango has seven distinctive vowel qualities. On the other hand, given that nasality is contrastive for five vowels, Sango has twelve phonemes of oral and nasal vowels. The reverse pattern of asymmetry is attested for Mauritian Creole, which, according to Baker (1972), has the five oral vowel phonemes /i, u, e, o, a/ and the three nasal phonemes /æ, 3, \tilde{a} /. All nasal vowels in this case have qualities distinct from the oral vowels, so that Mauritian Creole has eight vowel phonemes and eight distinctive vowel qualities.

We follow Maddieson's (1984: 128) practice to record vowel length contrasts as phonemic only if they are linked to vowel quality differences. In the present database this is applicable to Sri Lanka Creole Portuguese. This language has six vowels that occur as short and long, /i/, /u/, /e/, /o/, /o/ and /æ/. In addition, /ə/ appears only short, whereas /a:/ occurs only as a long vowel. Thus, in the present system Sri Lankan has eight vowel phonemes and eight distinctive vowel qualities.

According to Maddieson, it is often unclear in the description of the UPSID languages if the mid vowels, typically in five-vowel systems, are tense or lax. Consequently, he uses "e" and "o" to symbolize vowels in the mid range in such cases. Following this abstraction, I display mid vowels in Creole five-vowel systems as "e" and "o", even though the actual vowels may be tense or lax.

Creole phonology typology

The UPSID vowel quality inventories range in size from three to fifteen. Five-vowel systems are most frequent. The tripartite measure invoked for inventory size is carried over to the vowel qualities. Systems with three and four vowel qualities are classed as simple, inventories of five to seven qualities are classed as typical, whereas inventories of eight to fifteen vowel qualities are considered to be complex (see Maddieson 1984: 128).

No. of vowel qualities	No. of languages	% of languages
3	17	5.4%
4	27	8.5%
5	98	30.9%
6	60	18.9%
7	47	14.8%
8	17	5.4%
9	25	7.9%
10	15	4.7%
11	2	0.6%
12	5	1.6%
13	2	0.6%
14	0	0.0%
15	2	0.6%

 Table 3: Number of non-Creole vowel qualities (Maddieson 1984: 127)

Table 3 shows that around one-third of the UPSID languages exhibit inventories of five vowel qualities. Almost two thirds lie in the typical range of five to seven vowel qualities. Only slightly more than 20% show complex inventories of eight or more vowel qualities. The following displays show the vowel quality inventories of the Creoles in the sample.

(1) Five-vowel inventory

Bislama, Juba Arabic, Nubi, Kituba, Ndyuka, Palenquero, Tayo, Tok Pisin, Zamboangueño

(2) Six-vowel inventory(a) Baba Malay(b) Berbice

i			u	i	u	
	e		0	e	0	
		Э			3	
		а			а	

(3)	Seven-ve (a) Ange Sang	owel inve blar, Haiti o, São To	ntory an, Papiam omense, Sar	ento, amaccan	(b)) Negerhol	lands	
	i ε ε a	u o o			i	e E Ə a	u O	
(4)	Eight-vo (a) St. L	wel inver ucian	ntory (b) Krista	ang	(c) Mau	ritian	(d) Sri	Lankan
	i	u	i	u	i	u	i	u
	е	0	е	0	e	0	е	0
	ε	Э	e 3	С	æ	õ		e e
	аa		а		a	ã	æ	e a
(5)	Nine-vo Unserde	wel inven utsch	tory					
	i	u						
	1	U						
	е	0						

Perhaps most remarkable from the qualitative perspective are the six- and eight-vowel systems. Note that these systems mostly come about by the addition of one vowel to the five- or seven-vowel system, respectively. Baba Malay and Kristang have phonemic schwa in addition to the standard five and seven vowels, respectively. St. Lucian, Mauritian and Sri Lankan exhibit two vowels in the low range instead of one. A number of variations in the mid range of the vowel space may be observed. Berbice does not have a lax mid back vowel matching /ɛ/, whereas Sri Lankan has a schwa instead of /ɛ/. Similarly, the place of /ɔ/ is taken up by schwa in Negerhollands.⁴

12

ε

а

Э

The Negerhollands vowel inventory is taken from Sabino's (1990) discussion of the vowel system of the last speaker of the language.

Table 4 summarizes the number of creoles that show a given vowel quality system.

No. of vowel qualities	No. of Creoles	% of Creoles
5	9	39%
6	2	8.5%
7	7	30.5%
8	4	17.5%
9	1	4.5%

 Table 4: Number of Creole vowel qualities

The displays in (1) through (5) and Table 4 show that Creole languages are fairly narrow in the quantitative range of their vowel quality inventories. Creole languages span from five to nine vowels, unlike the three to fifteen distinct vowel qualities found in non-Creole languages. Five-vowel systems are most frequent in Creole languages, with seven-vowel systems running second. The great majority of Creole languages shows vowel quality inventories in the typical range of 5, 6 or 7 vowels, namely, 18 out of 23, that is 78%. The frequency of these systems in the Creoles is not significantly higher than that in the non-Creoles (205/317 = 65%) (Yates' chi-square (Preacher 2003): 1.205; p < 0.28).⁵ There are five complex systems of eight or more vowels, but no simple three- or four-vowel systems. The percentage of complex vowel systems is nearly identical for Creole (5/23 = 22%) and non-Creole languages (68/317 = 21%). The only difference approaching statistical significance is the robust presence of simple vowel inventories in the non-Creoles (44/317 = 14%) versus their absence in the Creoles (Fisher's exact: p < 0.056; two-tailed).

The Creole simplicity hypothesis predicts that there should be a substantial number of small and unmarked vowel quality inventories in Creole languages analogous to maximally unmarked /i, u, a/ in non-Creole languages like Classical Arabic, Miskitu or Australian Aboriginal languages. However, such inventories are absent from the Creoles in the sample. This lack of simple Creole vowel inventories is a significant problem for the Creole simplicity hypothesis.

One might ask if there are any reports of /i, u, a/ vowel inventories in Creole languages at all. In fact, I am aware of two systems that should be mentioned in this regard. Ehrhart (1993: 94) describes /i, u, a/ as the only vowels in a variety of Tayo she calls 'system A'. However, it is important to note that this system is described as moribund and in competition with the typical five-vowel system. The second report of an /i, u, a/ system comes from the Ngukurr dialect of Australian Kriol (Sandefur and Harris 1986: 180f.). However, it is significant that this system exists in variation with the standard five-vowel system, analogous to Tayo. The Creole simplicity hypothesis predicts that three-vowel systems should be common and stable in the Creole languages of the world. Instead, we find that such systems are exceedingly rare and unstable.

⁵ An anonymous reviewer has pointed out that the number of six-vowel systems in Creoles seems significantly smaller than that in non-Creoles. It is the task of future research to find an explanation for this discrepancy.

There is a good fit between the hypothesis that Creole phonologies are typical sound systems and the results of the present investigation of vowel quality inventories. Under this hypothesis, we expect to find a good majority of Creole languages with vowel quality inventory systems of five to seven members. By the same token, simple or complex vowel quality systems are not excluded *a priori*, but they are expected to play a subordinate role with respect to the middle-of-the-road systems. It is also expected under the Creole typicality hypothesis that the low and high ends of the range of the number of vowel qualities in non-Creole languages are not attested in Creole languages. In short, the numbers for vowel quality systems cluster around the center, not the simplicity end, in the Creole languages of the world. The investigation of stop consonants in the next section confirms this result.

5. Stop phonemes in Creoles

Stop consonants in the languages of the world appear in series in the sense that stops with different points of articulation such as labial, alveolar and velar appear as plain voiceless (/p, t, k/), plain voiced (/b, d, g/), aspirated voiceless (/p^h, t^h, k^h/) and so on. The number of stop series found in non-Creole languages ranges from one to six, as shown in Table 5.

No. of stop series	No. of languages	% of languages
1	50	15.8%
2	162	51.1%
3	76	24.0%
4	25	7.9%
5	2	0.6%
6	2	0.6%

 Table 5: Number of UPSID stop series (Maddieson 1984: 26)

Table 5 shows that more than half of non-Creole languages exhibit two series of stops and roughly a quarter display three stop series. Table 6 shows all fourteen (14) stop series attested in the UPSID database and their frequencies with respect to the overall sample.

Stop series	No. of languages	% of languages
Plain voiceless	291	91.8%
Plain voiced	212	66.9%
Aspirated voiceless	91	28.7%
Voiceless ejective	52	16.4%
Voiced implosive	35	11.0%
Prenasalised voiced	18	5.6%
Breathy voiced	7	2.2%
Laryngealised voiced	6	1.9%
Laryngealised voiceless	3	0.9%
Preaspirated voiceless	2	0.6%
Voiceless with breathy release	2	0.6%
Postnasalised voiced	1	0.3%
Prevoiced ejective	1	0.3%
Voiceless implosive	1	0.3%

 Table 6: Range and frequency of UPSID stop series (Maddieson 1984: 27)

Given the range, the number, and the frequency of stop series found in the UPSID languages, the complexity metric is as follows. A single stop series is considered simple. A language with one stop series almost invariably has plain voiceless plosives (49 out of 50 UPSID languages). Typical non-Creole languages have two or three series of stops. Plain voiceless stops are more frequent than plain voiced stops in languages with two series. Non-Creole languages with a third series typically add aspirated voiceless stops. Nearly three quarters of the UPSID languages have two or three series of stops. Languages with four, five, or six series of stops are complex. Less than 10% of non-Creole languages exhibit such complexity (see Maddieson 1984: 27ff.).

The series of stops found in the present sample of Creole languages are displayed in (6) through (10).

(6)	Plain voiceless	/p, t, k/	Baba Malay, Berbice, Bislama, Haitian,
	Plain voiced	/b, d, g/	Juba Arabic, Nubi, Kituba, Mauritian,
			Ndyuka, Negerhollands, Palenquero,
			Kristang, Papiamento, St. Lucian, Sri Lankan,
			Tok Pisin, Unserdeutsch, Zamboangueño
(7)	Plain voiceless	/p, t, k/	Тауо
	Prenasalized voiced	/ ^m b, ⁿ d, ⁿ g/	Ω.
(8)	Plain voiceless	/p, t, k/	Sango, Saramaccan
	Plain voiced	/b, d, g/	
	Prenasalized voiced	/"b. "d. "a/	

Thomas B. Klein

(9)	Plain voiceless Plain voiced Voiced implosive	/p, t, k/ /g/ /ɓ, ɗ/	São Tomense
(10)	Plain voiceless Plain voiced Voiced implosive Prenasalized voiced	/p, t, k/ /g/ /6/ / ⁿ b, ⁿ d, ⁿ g/	Angolar

All the languages in the right column of (6) have the two series named in the left column with the sounds listed in the center column; the meaning of the displays in (7) through (10) is analogous. Note that I take the presence of one or two stops in a series to be sufficient evidence to fully include it in the series count.

Table 7 summarizes the number of Creoles having a given number of stop series.

No. of stop series	No. of Creoles	% of Creoles
2	18	78%
3	4	17.5%
4	1	4.5%

 Table 7: Number of Creole stop series

Table 8 displays which stop series are attested in the present sample and their frequency in the Creoles.

	No. of Creoles	% of Creoles
Plain voiceless	23	100%
Plain voiced	22	95.5%
Prenasalized voiced	4	17.5%
Voiced implosive	2	9%

Table 8: Range and frequency of Creole stop series

The comparison of Tables 7 and 8 with Tables 5 and 6 shows that Creoles utilize a much narrower number and range of stop series than non-Creoles. No Creole language has been found to have a simple single stop series. The great majority of Creole languages (18 out of 23) have a prototypical dual series of plain voiceless and plain voiced stops. The remaining Creoles show substrate influence by exhibiting some of the distinct phonology of indigenous languages contributing to the language contact. Tayo does not have a plain voiced series, but boasts a prenasalized stop series instead. The Creoles with African substrates in (8) through (10) exhibit prenasalized voiced stops, voiced implosives, or both, an obvious African heritage.⁶ The occurrence of dual stops series is significantly higher in

16

⁶ An anonymous reviewer points out that Creoles with three stop series add either prenasalized voiced stops or voiced implosive stops to the most common series of plain voiceless and plain voiced stops. Phonemic aspirated voiceless stops, on the other hand, do not occur at all in the

Creole phonology typology

the Creoles than in the non-Creoles (Yates' chi-square (Preacher 2003): 5.305; p < 0.022). Only one Creole in the sample, Angolar, may be considered complex as far as the number and range of stop series found in it is concerned.

The Creole simplicity hypothesis predicts that there should be a substantial number of Creoles with a single, plain voiceless (/p, t, k/) stop series. However, none of the Creoles in the present database exhibits these stops as the only series. Instead, all Creoles in the sample have at least one more stop series. The absence of Creoles employing just a single stop series is unexpected and surprising under the Creole simplicity hypothesis and, hence, constitutes a significant problem for it. This begs the question if there are any reports of Creoles employing just a single /p, t, k/ stop series. In fact, Tryon (1987) describes Bislama as not having plain voiced stops. However, the minimal pairs in (11) obtained by Meyerhoff in her field work show that Bislama contrasts voiceless and voiced stops (personal communication 2001; see also Meyerhoff 2003).

(11) Dual stop series in Bislama

pig	ʻpig'	pen	'pen'
big	ʻbig'	ben	'bend'
traem	'try'	taon	'town'
draem	'dry'	daon	'down'
kol	'cold'	kad	'card'
gol	'goal, gold'	gad	'guard'

Given the data in (11), Bislama has the two standard series of stops.⁷ There appears to be no Creole with just one stop series.

Under the Creole typicality hypothesis, we expect to find a good majority of Creole languages with two and three stop series. Furthermore, it is expected that the great majority should have series of plain voiceless and plain voiced stops. Notwithstanding a small number of distinct systems, this is exactly what we find in the present database. Thus, the Creole typicality hypothesis makes the right quantitative predictions, whereas the Creole simplicity hypothesis does not.

present sample even though non-Creole languages with a third stop series typically add them. Given a possible connection to substrate inventories, the search for an explanation of this interesting discrepancy between Creoles and non-Creoles is beyond the scope of the present paper.

⁷ Unfortunately, the author did not have access to Crowley's (2004) new grammar of Bislama before finishing this chapter.

6. Conclusion

In this contribution I have measured Creole languages by what sounds they possess, not by which sounds may be absent. The typological approach taken is strictly synchronic and is based on a database of Creole languages from around the world. The present paper has focused on quantitative aspects of phonemic systems; segmental quality or markedness could not be treated in great detail. From this perspective, it must be concluded that the Creole simplicity hypothesis makes the wrong predictions. The segmental inventories of Creole languages are not simple, notwithstanding a very small number of exceptions. On the other hand, they do not tend to be very complex either. Instead, Creole inventories have a strong affinity to the typological middle.

When we look beyond segment numbers, it appears to be true that certain non-Creole languages may exhibit degrees of phonological complexity unmatched in Creoles. However, it is also the case that parts of the phonology or morphophonological patterns in Creoles have been discussed in recent work (Gooden 2003, James 2003, Klein 2003). Results from my typological investigation of Creole syllable structure (Klein 2004) show that the most frequent Creole syllable template is not just simple CV, as has been frequently claimed in the literature, but instead (C) (C) V (C). Furthermore, some Creoles exhibit significant syllabic complexity in that they allow complex codas (see also Plag and Schramm, this volume).

The results of the present study allow us to flesh out the idea that Creole languages are more alike than non-Creoles from the point of view of phonological segment inventories. In particular, Creole phoneme systems seem more alike because they occupy a typologically narrower range than non-Creole languages. Perhaps the most significant empirical result emerging from the current investigation is that the phoneme systems of Creoles are in the typical range compared to non-Creole languages when a consistent metric inspired by leading typological research is applied. Based on these results, I advance the Creole typicality hypothesis that Creole sound systems are of typical complexity and center on typological middle ground. This hypothesis can and should be tested in future research through typological investigations of markedness, additional segmental parameters, syllable structure and suprasegmentals.

References

Baker, Philip (1972): Kreol. A description of Mauritian Creole. London: Hurst.

Aceto, Michael (1996): Early Saramaccan syllable structure: An analysis of complex onsets from Schumann's 1778 manuscript. – In: Journal of Pidgin and Creole languages 11(1), 23-44.

Arends, Jacques, Pieter Muysken and Norval Smith (eds.) (1995): Pidgins and creoles. An introduction. Amsterdam: Benjamins.

- Bakker, Peter, Norval Smith and Tonjes Veenstra (1995). Saramaccan. In: Jacques Arends, Pieter Muysken and Norval Smith (eds.): Pidgins and creoles. An introduction, 165-178. Amsterdam: Benjamins.
- Baxter, Alan (1988): A grammar of Kristang (Malacca Creole Portuguese). In: Pacific Linguistics, Series B-95. Canberra: Australian National University.
- Bender, M. Lionel (1987): Some possible African Creoles: A pilot study. In: Glenn Gilbert (ed.): Pidgin and creole languages. Essays in memory of John E. Reinecke, 37-60. Honolulu: University of Hawai'i Press.
- Bickerton, Derek (1981): Roots of language. Ann Arbor: Karoma.
- Bickerton, Derek and Aquilas Escalante (1970): Palenquero: A Spanish-based creole of Northern Colombia. In: Lingua 24, 254-267.
- Carrington, Lawrence D. (1984): St. Lucian Creole. A descriptive analysis of its phonology and morpho-syntax. Hamburg: Buske.
- Corne, Chris (1999): From French to creole. The development of new vernaculars in the French colonial world. London: University of Westminster Press.
- Crowley, Terry (2004): Bislama reference grammar. Honolulu: University of Hawai'i Press.
- Ehrhart, Sabine (1993): Le créole français de St-Louis (le tayo) en Nouvelle-Calédonie. Hamburg: Buske.
- Fehderau, Harold (1962): Descriptive grammar of the Kituba language: A dialectal survey. Léopoldville: American Mennonite Brethren Board of Missions.
- (1966): The origin and development of Kituba (lingua franca Kikongo). Ph.D. dissertation. Cornell University.
- Forman, Michael (1972): Zamboangueño texts with grammatical analysis: A study of Philippine Creole Spanish. Ph.D. dissertation. Cornell University.
- Gooden, Shelome (2003): Prosodic contrast in Jamaican Creole reduplication. In: Ingo Plag (ed.): *Phonology and morphology of creole languages*, 193-208. Tübingen: Niemeyer.
- Holm, John (2000): An introduction to pidgins and creoles. Cambridge: Cambridge University Press. Huttar, George and Mary Huttar (1994): Ndyuka. London: Routledge.
- Institut National de Recherche et d'Action Pédagogiques (1982): Éléments de grammaire Kituba. Éditions Nathan.
- Ivens Ferraz, Luiz (1979): The creole of São Tomé. Johannesburg: Witwatersrand University Press.
- James, Winford (2003): The role of tone and rhyme structure in the organisation of grammatical morphemes in Tobagonian. – In: Ingo Plag (ed.): Phonology and morphology of creole languages, 165-192. Tübingen: Niemeyer.

Kaye, Alan and Mauro Tosco (2001): Pidgin and creole languages. A basic introduction. München: Lincom Europa.

- Klein, Thomas B. (2003): Syllable structure and lexical markedness in creole morphophonology: Determiner allomorphy in Haitian and elsewhere. – In: Ingo Plag (ed.): Phonology and morphology of creole languages, 209-228. Tübingen: Niemeyer.
- (2004): Diversity and complexity in the typology of syllables in creole languages. Ms. Georgia Southern University.
- Kouwenberg, Silvia (1994): A grammar of Berbice Dutch Creole. Berlin: Mouton de Gruyter.

Kouwenberg, Silvia and Eric Murray (1994): Papiamentu. München: Lincom Europa.

- Kouwenberg, Silvia and Pieter Muysken (1995): Papiamento. In: Jacques Arends, Pieter Muysken and Norval Smith (eds.): Pidgins and creoles. An introduction, 205-218. Amsterdam: Benjamins.
- Langsrud, Øyvind (2004): Fisher's exact test, Available from http://www.matforsk.no/ola/fisher.htm (last visited August 25, 2004).
- Laycock, Don (1984): Substratum elements in Tok Pisin phonology. In: Stephen A. Wurm and Peter Mühlhäusler (eds.): Handbook of Tok Pisin (New Guinea Pidgin), 295-307. Canberra: Australian National University.

- Lefebvre, Claire (1998): Creole genesis and the acquisition of grammar. The case of Haitian Creole. Cambridge: Cambridge University Press.
- Lewis, Anthony (1970): A descriptive analysis of the Palenquero dialect. M.A. thesis. Mona/ University of the West Indies.
- Lim, Sonny (1981): Baba Malay: The language of the 'straits-born' Chinese. M.A. thesis. Monash University.
- Lorenzino, Gerardo (1998): The Angolar Creole Portuguese of São Tomé. Its grammar and sociolinguistic history. München: Lincom Europa.

Lowry, Richard (2000): Fisher's exact probability test, Available from

http://faculty.vassar.edu/lowry/fisher.html (last visited August 23, 2004).

- Maddieson, Ian (1984): Patterns of sounds. Cambridge: Cambridge University Press.
- Maurer, Philippe (1995): L'angolar: Un créole afro-portugais parlé à São Tomé. Hamburg: Buske.
- (1998): El papiamentu de Curazao. In: Matthias Perl and Armin Schwegler (eds.): América negra, 139-218. Frankfurt: Vervuert.
- McWhorter, John (2001a): The world's simplest grammars are creole grammars. In: Linguistic Typology 5, 125-166.
- (2001b): What people ask David Gil and why: Rejoinder to the replies. In: Linguistic Typology 5, 388-412.
- Megenney, William (1986): El Palenquero: Un lenguaje post-criollo de Colombia. Bogotá: Instituto Caro y Cuervo.
- Meyerhoff, Miriam (2003): Reduplication in Bislama: An overview of phonological and semantic factors. In: Silvia Kouwenberg (ed.): Twice as meaningful: Reduplication in pidgins, creoles and other contact languages, 231-238. London: Battlebridge.
- Mufwene, Salikoko (1997): Kitúba. In: Sarah Thomason (ed.): Contact languages. A wider perspective, 173-208. Amsterdam: Benjamins.
- Mühlhäusler, Peter (1984): Internal development of Tok Pisin. In: Stephen A. Wurm and Peter Mühlhäusler (eds.): Handbook of Tok Pisin (New Guinea Pidgin), 75-166. Canberra: Australian National University.
- (1997): Pidgin and creole linguistics (2nd ed.). London: University of Westminster Press.
- Munteanu, Dan (1996): El papiamento, lengua criolla hispánica. Madrid: Gredos.
- Nida, Eugene (1956): Kituba grammar. Typescript. American Mennonite Brethren Mission.
- Owens, Jonathan (1997): Arabic-based pidgins and creoles. In: Sarah Thomason (ed.): Contact languages. A wider perspective, 125-172. Amsterdam: Benjamins.
- Pakir, Anne Geok-In Sim (1986): A linguistic investigation of Baba Malay. Ph.D. dissertation. University of Hawai'i.
- Parkvall, Mikael (2000): Out of Africa. African influences in Atlantic Creoles. London: Battlebridge.
- Pasch, Helma (1997): Sango. In: Sarah Thomason (ed.): Contact languages. A wider perspective, 209-270. Amsterdam: Benjamins.
- Pasch, Helma and Robin Thelwall (1987): Losses and innovations in Nubi. In: Philippe Maurer and Thomas Stolz (eds.): Varia Creolica, 91-165. Bochum: Brockmeyer.
- Patiño Rosselli, Carlos (1983): El habla en el Palenque de San Basilio. In: Nina S. de Friedemann and Carlos Patiño Rosselli (eds.): Lengua y sociedad en el Palenque de San Basilio. Bogotá: Instituto Caro y Cuervo.
- (1999): Aspectos de la estructura del criollo palenquero. In: Klaus Zimmermann (ed.): Lenguas criollas de base lexical español y portuguesa, 205-230. Frankfurt/Main: Vervuert.
- Plag, Ingo (ed.) (2003): Phonology and morphology of creole languages. Tübingen: Niemeyer.
- Plag, Ingo and Mareile Schramm (this volume): Early creole syllable structure: A cross-linguistic survey of the earliest attested varieties of Saramaccan, Sranan, St. Kitts and Jamaican, 131-150.

Preacher, Kristopher (2003): Interactive chi-square tests, Available from

http://www.unc.edu/~preacher/chisq/ chisq.htm (last visited August 25, 2004).

- Rountree, S. Catherine (1972): The phonological structure of stems in Saramaccan. In: Joseph Grimes (ed.): Languages of the Guianas, 22-27. Norman: SIL.
- Sabino, Robin (1990): Towards a phonology of Negerhollands: An analysis of phonological variation. Ph.D. dissertation. University of Pennsylvania.
- Sandefur, John and John Harris (1986): Variation in Australian Kriol. In: Joshua A. Fishman et al. (eds.): The Fergusonian impact. Vol. 2: Sociolinguistics and the sociology of language, 179-190. Berlin: Mouton de Gruyter.
- Singh, Rajendra and Pieter Muysken (1995): Wanted: A debate in pidgin/creole phonology. In: Journal of pidgin and creole languages 10, 157-169.
- Smith, Geoff (2002): Growing up with Tok Pisin. London: Battlebridge.
- Smith, Ian Russell (1977): Sri Lanka Creole Portuguese phonology. Ph.D. dissertation. Cornell University.
- Stolz, Thomas (1986): Gibt es das kreolische Sprachwandelmodell? Frankfurt: Lang.
- Swift, Lloyd B. and Emile W. A. Zola (1963): Kituba basic course. Washington, D.C.: Foreign Service Institute.
- Thomason, Sarah (ed.) (1997): Contact languages. A wider perspective. Amsterdam: Benjamins.
- Tinelli, Henri (1981): Creole phonology. The Hague: Mouton.
- Tryon, Darrell (1987): Bislama. An introduction to the national language of Vanuatu. In: *Pacific Linguistics*, Series D-72. Canberra: Australian National University.
- Verhaar, John (1995): Toward a reference grammar of Tok Pisin. An experiment in corpus linguistics. Honolulu: University of Hawai'i Press.
- Volker, Craig (1982): An introduction to Rabaul Creole German (Unserdeutsch). M.A. thesis. University of Queensland.
- Walker, James and William Samarin (1997): Sango phonology. In: Alan Kaye (ed.): Phonologies of Asia and Africa (including the Caucasus), Vol.2, 861-880. Winona Lake: Eisenbrauns.
- Whinnom, Keith (1956): Spanish contact vernaculars in the Philippine Islands. Hong Kong: Hong Kong University Press.
- Wurm, Stephen A. and Peter Mühlhäusler (eds.) (1984): Handbook of Tok Pisin (New Guinea Pidgin). In: *Pacific Linguistics*, Series C-70. Canberra: Australian National University.

Thomas B. Klein, Georgia Southern University, Department of Writing and Linguistics, P.O. Box 8026, Statesboro, GA 30460, USA, Email: tklein@georgiasouthern.edu

