DGfS-CNRS Summer School on Linguistic Typology
Leipzig 15 August–3 September 2010

Creole Languages in a World-Wide Perspective

SUSANNE MICHAELIS

Max-Planck-Institut für evolutionäre Anthropologie & Universität Gießen
Class 8, 3 September 2010

1. How complex are pidgin & creole languages compared with other languages?

• notion of complexity in creoles has been very prominent over the last 15 years (as also within the field of general linguistics; see Sampson & Gil & Trudgill (2009) (eds.), *Language complexity as an evolving variable*), but the definitions of "complexity" vary widely.

• 'complexity' may be understood as 'processing complexity or difficulty', but this notion of complexity is left aside here

• here, we are looking at **'structural complexity'**: e.g. the more distinctions there are in a paradigm, the more complexity there is

2. Do creoles have the simplest grammars?

• McWhorter (1998, 2001a): creoles have the simplest grammars within the world's languages:

  (a) little or no inflectional morphology,
  (b) no lexical or morphosyntactic tone,
  (c) no non-transparent derivational morphology;

• during the pidgin phase all 'ornamental' marking has been lost, creoles are too young to have been able to develop complex features of older languages

3. How complex are pidgin & creole languages compared with other languages? A partial replication of Parkvall (2008)

3.1 Parkvall 2008

• Parkvall examines 53 *WALS* features which reflect degrees of complexity, and codes them additionally in 29 pidgins and creoles, comparing the degree of complexity of these creoles with 153 non-creoles of the *WALS* sample.
• We used the following 18 features (a subset of Parkvall's 53 features) to measure complexity:

13   Gender Distinctions in Independent Personal Pronouns
15   Inclusive/Exclusive Distinction in Independent Personal Pronouns
18   Politeness Distinctions in Second Person Pronouns
22   Occurrence of Nominal Plural Markers
28   Definite Articles
29   Indefinite Articles
33   Distance Contrasts in Demonstratives
36   Numeral Classifiers
54   Suppletion according to Tense and Aspect
58   Alignment of Case Marking of Full Noun Phrases
59   Alignment of Case Marking of Pronouns
70   Comitatives and Instrumentals

71 Noun Phrase Conjunction and Comitative
72 Nominal and Verbal Conjunction
73 Predicative Noun Phrases
76 Predicative Noun Phrases and Predicative Locative Phrases
91 Applicative Constructions
120 Tone

Feature 22: Occurrence of Nominal Plural Markers

Feature 28: Definite Articles

Feature 54: Suppletion according to Tense and Aspect

Feature 120 Tone

3.2 Parkvall's ranking of *WALS* languages (some creoles included)

• Pidgins and creoles which Parkvall added to the *WALS* sample:

rank	language	complexity score (between 0 and 1)			
1.	Russenorsk	0.00	…		
2.	Lingua Franca	0.06	18.	Principense	0.23
			…		
			24.	Saramaccan	0.25
3.	Bislama	0.13	…		
4.	Chinuk Wawa	0.15	28.	Haitian Creole	0.29
…			29.	Santome	0.31
10.	Mauritian Creole	0.20	30.	Papiamentu	0.33

• thus Parkvall concludes: " Typologically speaking, creoles stand out from languages in general, and the most salient difference is that they present a **lower structural complexity**. This does not necessarily have any bearing on issues regarding psycholinguistic complexity, however, and certainly not on their expressive potential." (Parkvall 2008: 283, my emphasis)

• average complexity 0.22, i.e. like Ndyuka in *WALS*

3.3 Some initial results from *APiCS*

• We took a preliminary look at 18 *APiCS* features which are also found in *WALS*, a subset of Parkvall's 53 features. We computed similar complexity scores for our 68 *APiCS* languages, as well as for 75 *WALS* languages.

• We have much fewer features than Parkvall, but more pidgin/creole languages, and our data are more complete.

- Results: average complexity value:

WALS languages	0.45
APiCS languages	0.38

- The results for the *WALS* languages are very similar to Parkvall's results:
  - average complexity score is 0.45 (compared to Parkvall's 0.42)
  - the languages are ranked in a very similar way (rank correlation test: Kendall's tau 0.51, p = 2.2 * 10^{-16}; statistical method whether two rankings are significantly similar)

- However, **the results for pidgin/creole languages are quite different**: In the *APiCS* data, pidgins/creoles are much more complex, and much less distinct from *WALS* languages. Why?

- One possible answer: The 18 features that we considered happen to be features with respect to which the pidgin/creole languages are not significantly simpler.

- This leads to the question:
Which features account for the simplicity that Parkvall discovered?

• Such research questions can be answered using the *APiCS* database.

• Another possible answer: Parkvall's pidgin/creole sample embraces 30 languages, 14 with West African substrate (Atlantic and Caribbean creoles), 4 with Pacific languages, 2 with East African languages (Indian Ocean), 1 with South-East Asian substrate, and some others (North/South America)

--> totally West African/Niger-Congo biaised sample

• we have already seen the **typologically different behavior of creoles of India/SE Asia/Pacific** regarding other features compared to the Atlantic and African creoles (e.g. Alignment of nouns and pronouns, inclusive/exclusive distinction, Unmarked stative and dynamic verbs)

• if e.g. Ternate Chabacano, Diu Indo-Portuguese, and Sri Lankan Portuguese with their complexity scores would have been the prototypes for creole languages and if they would have been as well studied as the Caribbean creoles, our whole picture of creoles, and especially the far-reaching hypotheses about innateness of creole grammars, simplicity of creole grammars etc. would have

been very different.

• John Holm (2008) focusses on the different results of language contact for the Atlantic creoles compared with the Asian creoles:

> The reason for the widespread belief that creolization entails the loss of all inflectional morphology grew out of the general absence of inflectional morphemes as such in the group of creoles that has been most studies, the Atlantic creoles of the Caribbean region and coastal West Africa. All of these creoles have the Western European lexifiers that were the languages of empires (Portuguese, Spanish, French, Dutch and English) and Niger-Congo or West African substrates, i.e. at least partially inflected superstrates and largely non-inflected, isolating substrate languages. And while the coming together of such languages under the conditions that produce creoles was indeed unlikely to produce creoles that retain any inflections, the complete loss of inflectional morphology is not an inherent part of the process of creolization. All that was missing to disprove this assumption was a careful study of a creole that grew out of a superstrate and a substrate that are both inflected. (Holm 2008: 319f.)

• compare e.g. the Indo-Portuguese creole Korlai which has the following verbal inflectional morphology: *katad* 'sung' (< P *cantado*), *katan* 'is singing' (< P

cantando), *kato* 'sang' (< P *cantou*), see also Diu, and Sri Lanka Portuguese

• *APiCS* comes at the right time for a thorough revision of what we think creole languages are.

• urgent need for sampling for genealogical control within contact languages

4. Sampling for genealogical control within contact languages

• sampling for genealogical control is a very much debated question within typology (cf. Bickel 2008), but not yet in comparative contact linguistics (neither in SLA nor in creoles/pidgins, see first weak attempts in Kouwenberg 2010a and 2010b)

• the question of the genealogical affiliation of creoles and pidgins has been controversially discussed

• regardless of one's theory of creolization (continuation of European varieties vs. new contact languages via abrupt creolization vs. relexification etc.), it is clear that

in many creolization situations the main lexical input comes from the (European) base language(s) whereas structural grammatical features can be traced back (to different degrees) to the relevant substrate language(s).

• "genealogically" controlled sample of **contributing pair types** in contact languages

5. Contributing pair types

Some contributing pair types from *APiCS*:

Niger-Congo	<--->	European (main Caribbean creoles)
Ijoid	<--->	European (Berbice Dutch)
Malagasy/Eastern Bantu	<--->	European (French Indian Ocean Creoles)
Indic	<--->	European (Diu, Korlai)
Oceanic	<--->	European (Bislama, Tok Pisin, Tayo)
Papuan	<--->	Papuan (Pidgin Yimas-Arafundi)
Australian	<--->	European (Kriol)

Nilotic	<--->	Arabic (Kinubi, Juba Arabic)
Bantu	<--->	Bantu (Kikongo-Kituba, Lingala)
	<--->	European (Singlish, CPE, CRP)
Central Malayo-Polynesian	<--->	Malay (Ambon Malay)
Western Malayo-Polynesian	<--->	European (Cavite, Ternate, Zamboanga Chabacano)

• for languages resulting from one and the same contributing pair type (e.g. Caribbean creoles), only **one** language should be chosen for the genealogically controlled set of languages, **unless there is variation in the feature values**;

• e.g. if all sampled Caribbean creoles show the same value "Double-Object construction" in the Feature "Ditransitive constructions", then this should be counted only as ONE instance of this feature value compared to SE Asian languages where there are "Indirect-Object constructions" and "Primary-Object constructions".

• calculations should be based on contributing pair types and not on indidual languages in order to get a more realistic view on "typical" creole features

6. Two main claims within creole studies which we disconfirmed

(i) creole languages are to a large extent uniform, all exhibit a range of "typical" creole features (Bickerton 1981ff.; McWhorter 1998ff.)

• This is not true: the *APiCS* data clearly showed that there is an impressive diversity within creoles (and other contact languages) for many structural features. There are most importantly **linguistic areas**, Atlantic & Caribbean creoles vs. India & Asia/Australia. This is clearly due to similar language contact outcomes from similar contributing language pairs.

(ii) creole languages show the "world's simplest grammars" (McWhorter 1998ff.)

• this claim was already disconfirmed before the *APiCS* data have been available (e.g. Gil 2001 on Riau Indonesian). But the replication of Parkvall 2008 with a subpart of the *APiCS* data also showed counter-evidence to this "simplicity claim". Further detailed studies on the wealth of the *APiCS* data will give us new insights on the nature of creoles.

APiCS/*WALS* complexity ranking based on 18 features

Language		Language		Language	
Ternateño	.86	Alamblak	.57	Creole (São Vicente)	
Khoekhoe	.73	Japanese	.57	Guinea Bissau Creole	.50
Spanish	.72	Kanuri	.57	Sri Lanka Malay	.50
Meithei	.67	Korean	.56	French	.50
Abkhaz	.63	Krongo	.56	Georgian	.50
Louisiana Creole	.63	Maori	.56	German	.50
Sri Lanka Portuguese	.62	Supyire	.56	Tagalog	.50
Greek (Modern)	.62	Zamboangueño	.56	Hungarian	.49
Ju‖'hoan	.62	Diu Indo-Portuguese	.55	Persian	.48
Kannada	.62	Guyanese Creole English	.54	Afrikaans	.48
Ainu	.61	Lezgian	.53	Ghanaian Pidgin	.48
Burushaski	.60	Yaqui	.53	Fijian	.47
Malagasy	.60	Michif	.52	Hindi	.47
Basque	.59	Hausa	.51	Khasi	.47
Mandarin	.59	Latvian	.51	Lakhota	.47
		Cape Verdean	.51	Mangarrayi	.47

Language		Language		Language	
Nicaraguan Creole	.47	Vincentian Creole	.44	Trinidadian Creole	.41
Saramaccan	.47	English	.43	Acoma	.40
Hebrew (Modern)	.46	Guaraní	.43	Finnish	.40
Turkish	.46	Jakaltek	.43	Imonda	.40
Chamorro	.45	Russian	.43	Gurindji Kriol	.40
Yukaghir (Kolyma)	.45	Yoruba	.43	Hawaiian	.40
Bahamian Creole	.45	African American Vernacular English	.43	Creole English	
Gullah	.45	Lingála	.43	Media Lengua	.40
Korlai Creole Portuguese	.45	Singlish	.43	Nigerian English-lexifier Pidgin/Creole	.40
Pichi	.45	Sranan	.43	Slave	.39
Oromo (Harar)	.44	Burmese	.42	Swahili	.39
Belizean Creole	.44	Chukchi	.42	Greenlandic (West)	.38
Cape Verdean Creole (Santiago)	.44	Cape Verdean Creole (Brava)	.41	Thai	.38
		Sierra Leone Krio	.41	Mbugu	.38
				San Andrés Creole	.38

Yidiny .37
Batavia Creole .37
Ngiyambaa .36
Vietnamese .36
Bislama .36
Chinese Pidgin .36
English
Lungwa .36
Santome
Nengee .36
Roper River .36
Creole
Ewe .35
Iraqw .34
Khalkha .34
Tiwi .34
Zulu .34
Cameroon .34
Pidgin English
Amele .33

Lango .33
Palenquero .33
Reunion Creole .33
Yimas- .33
Arafundi
Pidgin
Indonesian .32
Evenki .31
Hunzib .31
Mapudungun .31
Quechua .31
(Imbabura)
Angolar .31
Bazaar Malay .31
Berbice Dutch .31
Creole
Early Sranan .31
Principense .31
Haitian Creole .30
Arabic .29

(Egyptian)
Hixkaryana .29
Papia Kristang .29
Ket .28
Kobon .28
Fa d'Ambu .27
Kinubi .27
Mauritian .27
Creole
Sango .27
Chinese .25
Russian Pudgin
Jamaican .25
Creole
Negerhollands .25
Seychelles .25
Creole
Tok Pisin .25
Yagua .25
Pidgin .23

Hindustani		Guadeloupe Creole	.16	Ambon Malay	.15
Chinuk Wawa	.22			Eskimo Pidgin	.01
Juba Arabic	.20	Martinique Creole	.16		
Canela-Krahô	.18				
Pirahã	.18	Tayo	.16		

References

Bickel, Balthasar. 2008. A refined sampling procedure for genealogical control. In: *STUF* 2008:61 (3), 221-233.

Gil, David. 2001. Creoles, Complexity and Riau Indonesian. *Linguistic Typology* 5: 325–371.

Holm, John. 2008. Creolization and the fate of inflections. In: Stolz, Thomas & Bakker, Dik & Salas Paloma, Rosa (eds.), *Aspects of Language Contact. New Theoretical, Methodological and Empirical Findings with Special Focus on Romancisation Processes*. Berlin: Mouton, 299-324.

Kouwenberg, Silvia. 2010a. Creole studies and linguistic typology. Part 1. In:, *JPCL* 25.1, 173-186.

Kouwenberg, Silvia. 2010b. Creole studies and linguistic typology. Part 2. In:, *JPCL* 25.2, 359-380.

McWhorter, John. 2001a. The world's simplest grammars are creole grammars. In: *Linguistic Typology* 5, 125-166.

McWhorter, John. 2001b. What people ask David Gil and why: Rejoinder to the replies. In: *Linguistic Typology* 5, 388-412.

McWhorter, John. 1998. Identifying the creole prototype: Vindicating a typological class" In: *Language* 74 (4): 788–818.

Parkvall, Mikael. 2008. The simplicity of creoles in a cross-linguistic perspective. In: Miestamo, Matii, Sinnemäki, Kaius & Karlsson, Fred (eds.), *Language complexity*: *Typology, contact, change*. Amsterdam: Benjamins, 265-285.

Sampson, Geoffrey & Gil, David & Trudgill, Peter. 2009. (eds.), *Language complexity as an evolving variable*, Oxford: OUP.