World-wide comparative evidence for calquing of valency patterns in creoles

SUSANNE MARIA MICHAELIS
Leipzig University &
Max Planck Institute for the Science of Human History (Jena)

1. Introduction

Creole languages mainly evolved in the context of European colonial expansion during the 16th through 19th centuries where the European colonial language, the lexifier language, mixed with different labor-force languages spoken by the slaves or indentured laborers, the substrate languages. Creole languages overwhelmingly show lexical material from the European (or other-based) lexifier languages. In Seychelles Creole, for instance, over 90% of the verbs stem from dialectal 18th century French varieties: *donnen* 'give' < French *donner, koze* 'speak' < French *causer,* and *gete* 'watch' < French *guetter*, *bezwen* 'need' < French (*avoir*) *besoin* (*de*). But creole languages consistently show valency patterns that cannot be traced back to their lexifier languages, but derive from their substrate languages.

Substrate influence in creoles has been discussed by various authors (e.g. Alleyne 1980, Boretzky 1983, Holm 1988, Lefebvre 1998, 2011, 2015, McWhorter 1997; Parkvall 2000, Siegel 1999, 2008; Winford & Migge 2007) but they have focussed on specific constructions, on specific creole languages and/or specific substrate languages. Such claims are often plausible but they do not convince all creolists as there are also plausible alternative explanations in terms of the lexifier or universal tendencies. Claims of substrate effects are most convincing when the substrate feature is very rare cross-linguistically (see Parkvall 2000) and does not occur in the lexifier. A well-known example is the trial number in Tok Pisin personal pronouns (*yumitripela* 'the three of us'), which was calqued on an Oceanic substrate language (Smith & Siegel 2013b: 217).

In this paper, I start out from the observation that a convincing case for substrate influence can also be made by adopting a world-wide comparative approach. If there are recurrent matches between substrate and creole structures in a given construction type, in creoles of different world-regions and with different substrates, then we can exclude the possiblity of an accident, and substrate influence is the only explanation.[1]

We will see that in the domain of valency, creoles show a substantial amount of constructional calquing from their substrates. The construction types that I will look at are ditransitive constructions (§3), weather constructions (§4), experiencer constructions (§5), and motion constructions (§6).

I will draw on typological data from the *Atlas of Pidgin and Creole Language Structures* (Michaelis et al. 2013a, 2013b), the *World Atlas of Language Structures* (Haspelmath et al. 2005; Dryer & Haspelmath 2013), and other typologically relevant data where necessary. My conclusion is that the data provided in *APiCS* support the claim that during creolization, valency patterns have been systematically calqued into the nascent creoles.

In each section, I will contrast the creole patterns with both the lexifier and substrate patterns to demonstrate how closely the creoles imitate the valency patterns of their substrates (or adstrates) instead of continuing the patterns of their lexifier languages. The analyzed data allow an even more far-reaching claim, namely that the creators of creole languages have not only imitated some valency patterns, but the semantic construal patterns of their dominant (native) languages. A European-based creole construction 'Rain falls' may seem to be not that dramatically different from the canonical European lexifier construction '(It) rains'. But the fact that in the creole the weather event is coded in subject position combined with some general verb whereas in the European languages the event is coded in the verb with an expletive subject pronoun (in English: ***it*** *rains*, and French: ***il*** *pleut*), illustrates that the whole framing of the weather event is crucially different in both sets of languages. Some creolists may prefer to explain such a difference in construction in terms of transparency and/or (semantic/syntactic) simplicity. But a closer and typologically informed look at the data

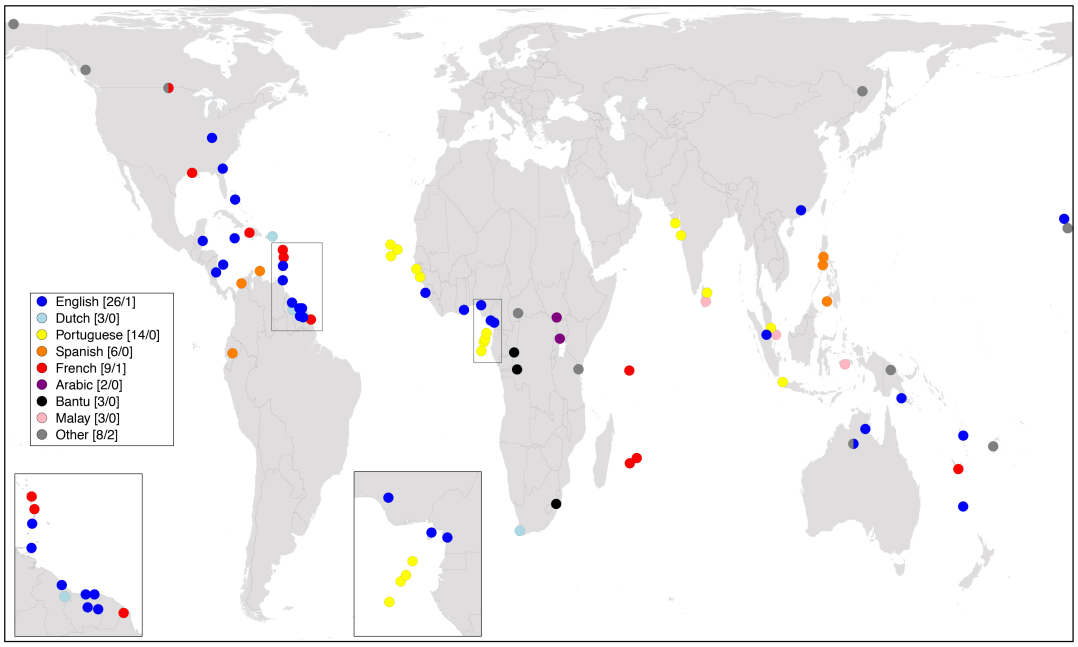
[1] A study that makes this argument is Michaelis & Haspelmath 2003 on ditransitive constructions, and McWhorter (1997: 35-39) similarly argues for substrate influence in a wide range of serial verb constructions for Atlantic and non-Atlantic creoles.

will uncover that the reason for the creole languages to show these valency patterns is the result of systematic constructional calquing of the substrate/adstrate patterns.

Before I discuss the four different construction types, I will brievely present the source of the creole data used in this study.

2. The *APiCS* database

The present paper is based on the *Atlas of Pidgin and Creole Language Structures* (*APiCS*, Michaelis et al. 2013a, 2013b), a large-scale comparative database of pidgins, creoles and mixed languages. In *APiCS*, 76 contact languages world-wide are investigated with respect to 130 structural features, some 330 segmental features, and 28 sociolinguistic features. As Map 1 illustrates, *APiCS* covers all major world regions. Compared to other existing cross-creole studies (e.g. Holm & Patrick 2007, Parkvall 2008), *APiCS* comprises information on a substantial number of non-Atlantic contact languages in South Asia, Southeast Asia, Melanesia, and Australia. It also contains contact languages which have non-European base languages, like Arabic, Bantu, and Malay.



Map 1: 76 contact languages in *APiCS*[2]

[2] For space reasons, the labels of the languages cannot be given in the map. For a list of the languages, see *APiCS Online* http://apics-online.info/contributions#2/30.3/10.0. Mixed languages are represented by a

In contrast to the *World Atlas of Language Structures* (*WALS*, Haspelmath et al. 2005; Dryer & Haspelmath 2013), *APiCS* follows an expert-based typology-approach[3]: Each language expert or team of experts has filled in a structure questionnaire designed by the four editors (Michaelis, Maurer, Haspelmath, and Huber) by choosing feature values and giving examples which illustrate the value assignment. *APiCS* has been published as a four volume print work: the first volume displays chapters written by the four editors on each of the 130 structural features with the corresponding map where each coloured dot represents the given feature value for a given language, and the three other volumes constitute the *Survey of pidgin and creole languages* where each expert gives a concise overview of the social history, sociolinguistic situation and the grammar of the contact language. Besides the paper publication, the underlying database *APiCS Online* has been published electronically with open access (apics-online.info), with more than 15,000 fully glossed and translated examples as well as many references. 48 structural features of *APiCS* have been taken over from *WALS*. This gives us the unique possibility to see the *APiCS* languages against the background of languages world-wide. *APiCS Online* provides special views for this comparison[4].

It is noteworthy that creolists have made only little use of the typological wealth presented in *WALS* (but see Bakker et al. 2011, Daval-Markussen 2014 on quantitative comparisons). This may be due to different reasons: The majority of creolists are language specialists, and only a small minority of them is interested in broad language comparison. Furthermore, the discussion has been centered around the question whether mainly (or solely) substrates or lexifiers or universals have shaped creoles (see e.g. Muysken & Smith 1986, Lefebvre 1998, Chaudenson 1991, Mufwene 2001). But much less attention has been paid to a solid *qualitative* comparison of historically unrelated creoles with their substrates and their lexifiers to eventually filter out *which kind of features* can be systematically traced back to substrates or lexifiers or universal processes during creolization. I think that with *WALS* and other typological studies at

---

pie-chart of the two contributing languages, e.g. Michif, a mixed language of Northern America, has half a red dot for the French component, and half a grey dot for other, in this case Cree.

[3] See also other CLLD-database publications with this same approach (http://clld.org/datasets.html), *The World Loanword Database, The Electronic World Atlas of Varieties of English.*

[4] All *APiCS* maps in this paper were designed by Hans-Jörg Bibiko, Max Planck Institute for the Science of Human History (Jena).

hand, and the publication of *APiCS* a new round of meticulous qualitative – and not only quantitative – language comparison has been opened up.

This article is meant to be a starting point for such a qualitative, data-based re-evaluation of the role of the contributing languages and universal features in the process of creolization.

In my view, every theory of pidginization and creolization should base itself on such a fine-grained data-driven study on the world-wide variation of creole languages and non-creole languages to make any significant claims about creole languages in general. Even more so if far-reaching universal and/or cognitive processes are invoked.

In the following sections, I will discuss four construction types in the 59 creoles in *APiCS*. I use the term "creole" for classical creoles like Saramaccan, Mauritian Creole, and Tayo, but also pidgincreoles, i.e. expanded pidgins that are used in a wider set of linguistic functions even though they are not the mother tongues of all their speakers, such as Cameroon Pidgin English or Tok Pisin (Bakker 2008: 131ff, who takes this term from Philip Baker).

As for the cross-linguistic data available for the four construction types, it is only ditransitive constructions with the verb 'give' (§3) that are covered by *WALS*. But weather constructions with rain (§4) have been analysed by a series of typological articles, e.g. by Eriksen et al. 2010, 2012. For motion-to/motion-from constructions (§4), we have an excellent typological study by Wälchli & Zuñiga 2006 which I will rely on. Unfortunately, there are no large-scale typological surveys for experiencer constructions (§6). Nevertheless, I have tried to collect the available typological information for the construction under discussion.

3. Ditransitive constructions

The first construction type that I will discuss in this paper is DITRANSITIVE CONSTRUCTIONS. Such constructions involve a verb of physical or mental transfer with three participants: an agent, a recipient, and a theme. Following the *APiCS* chapter (Haspelmath, Michaelis & *APiCS* Consortium 2013), I will only look at the most frequent physical transfer verb 'give', as constructions with other ditransitive verbs may show considerable variation. Furthermore word order will not be considered. Examples

(1) and (2) feature a double-object construction, with no preposition marking either recipient or theme (corresponding to the same coding of the patient in a monotransitive clause).

(1) Krio (English-based; Finney 2013)

di	*uman*	*gi*	*di*	*titi*	*sɔm*	*mɔni*
the	woman	give	the	girl	some	money
			REC		THEME	

'The woman gave the girl some money.'

(2) Seychelles Creole (French-based; Michaelis & Rosalie 2013)

Mon	*'n*	*donn*	*Marcel*	*en*	*mang.*
1SG	PRF	give	Marcel	a	mango
			REC		THEME

'I gave Marcel a mango.'

Examples (3) and (4) show an indirect-object construction where the recipient is marked by a special preposition (against the identical marking of theme and patient in an monotransitive clause), *ku* in Papiá Kristang and *long* in Tok Pisin.

(3) Papiá Kristang (Portuguese-based; Baxter 2013)

eli	*sa*	*tiu*	*ja*	*bendé*	*aké*	*prau*	***ku***	*yo*
3SG	GEN	uncle	PFV	sell	that	boat	DAT	1SG
					THEME			REC

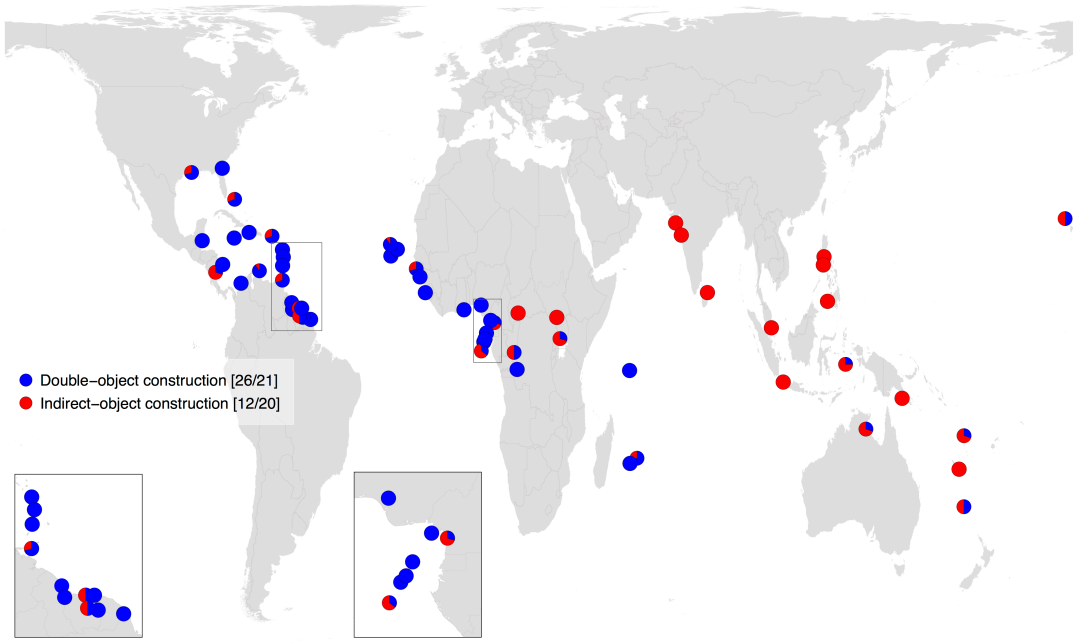
'His uncle sold the boat to me.'

(4) Tok Pisin (English-based; Smith & Siegel 2013a)

Givim	*mani*	***long***	*papa*	*bilong*	*yu.*
Give	money	PREP	father	POSS	2SG
	THEME		REC		

'Give your father the money.'

It has been claimed that creoles typically show double-object constructions (Bickerton 1995:1453, Bruyn et al. 1999) even if their lexifiers, for instance the Romance languages, have an indirect-object construction (e.g. French *J'ai donné une mangue* ***à*** *Marcel*. 'I gave Marcel a mango'). But if we look at the world-wide distribution of creoles in *APiCS*, the picture is not uniform at all (Map 2).



Map 2: Ditransitive constructions with 'give' in 59 creoles of *APiCS* (Haspelmath, Michaelis & *APiCS* Consortium 2013); the numbers in brackets refer to the number of creoles which show the feature value in question, e.g. 26 creoles in *APiCS* exclusively have double-object constructions, and 21 creoles share this construction with the indirect-object construction, the pie-charts reflecting the relative importance of each feature value.

First of all, creole languages can have both construction types, double-object constructions (blue dots) and indirect-object constructions (red dots)[5], which is shown in the pie-charts on Map 2. But if for simplicity we restrict ourselves to creoles with

[5] We do not investigate the secondary-object construction due to its marginal status in the contact languages that are studied here.

exclusive double-object constructions and exclusive indirect-object constructions, that is representing single-coloured dots, the figures are as follows[6]:

Table 1: Ditransitive constructions in creoles (exclusive marking only)

	APiCS creoles		*APiCS* bi-clans of creoles	
Double-object constructions	26	69%	9	**56%**
Indirect-object constructions	12	31%	7	**44%**
Total	38	100%	16	100%

Indeed a clear majority of *APiCS* creoles (69%) feature the double-object construction whereas only a third of *APiCS* creoles show the indirect-object construction as in (3-4). But I have given yet another figure in Table 1, the so-called BI-CLAN distribution. What do I mean by this?

Each creole evolved historically from at least two contributing languages, a lexifier and a substrate (or a set of substrates). For instance, English-based Atlantic creoles, i.e. languages like Jamaican, Trinidad English Creole, Sranan, and Krio are historically much more closely related to each other than, for instance, Jamaican is related to Tok Pisin. The former all stem from an encounter of English dialectal varieties of the 17th/18th centuries and African languages of the so-called Macro-Sudan Belt (see Güldemann 2010)[7]. Tok Pisin, too, is an English-based creole, but compared to the Atlantic creoles just mentioned it has very different substrate/adstrate languages, namely Oceanic and Papuan languages of Melanesia. In order to assess these different historical profiles of the various creoles and to control for genealogical and areal relatedness of both the substrate(s) and the lexifier, I propose a new sampling method,

[6] For the full picture, see Michaelis & Haspelmath (2003) and Haspelmath, Michaelis and the *APiCS* Consortium (2013).

[7] The situation is actually more complex than this because there are sometimes specific historical links between varieties within a bi-clan, for instance where a new island was settled from a creole speaking island (cf. Seychelles Creole which is an offspring of Mauritian Creole). Cf. also Huber 1999 on the settlement of Freetown, Sierra Leone, from the New World (Nova Scotian settlers and Jamaican Maroons).

which I call BI-CLAN SAMPLING. A CLAN[8] is a language or a family or a linguistic area, and a BI-CLAN is a combination of a lexifier clan and a substrate/adstrate clan[9]. For example, the lexifier clan "English" combined with the substrate clan "Macro-Sudan" gives rise to the bi-clan "English/Macro-Sudan" which comprises creoles such as Jamaican, Sranan, Saramaccan, Trinidad Creole English, Gullah, Bahamian, San Andrés Creole etc. The need for such a new sampling method in creole studies has become obvious because all available current samples of creole languages (Holm & Patrick 2007, Parkvall 2008, *eWAVE* 2011, *APiCS* 2013) are heavily biased towards one major historical group of creoles, namely Atlantic creoles, i.e. creoles spoken in West Africa and the Caribbean. These creoles have arisen via the contact between some very closely related Western European lexifier languages, namely English, Dutch, French, Portuguese, and Spanish, and a group of West African Niger-Congo substrate languages, which stretch over several different language families, but nevertheless due to longstanding language contact, show clear patterns of structural convergence (see Güldemann 2010). Given the numerical overrepresentation of one group of historically closely related creoles, i.e. Atlantic or more precisely European/Macro-Sudan creoles, the bi-clan approach then reduces the genealogical and areal predominance of Atlantic creoles in a principled way to account for the world-wide diversity of creoles. Such an approach is necessary if we want to say something about creoles *in general*, and not just about one set of historically closely related creoles.

With this explication of bi-clan in the background, we can go back to the distribution of the figures for double-object and indirect-object constructions in Table 1.

How do we determine the bi-clan numbers for the ditransitive constructions in Table 1? Creole languages of the same bi-clan are genealogically (and/or areally) closely related, so they are likely to show similar typological profiles. This means that languages of the same bi-clan often show the same feature value for a given feature. In the present

[8] The term "clan" was suggested to me by Bernard Comrie.

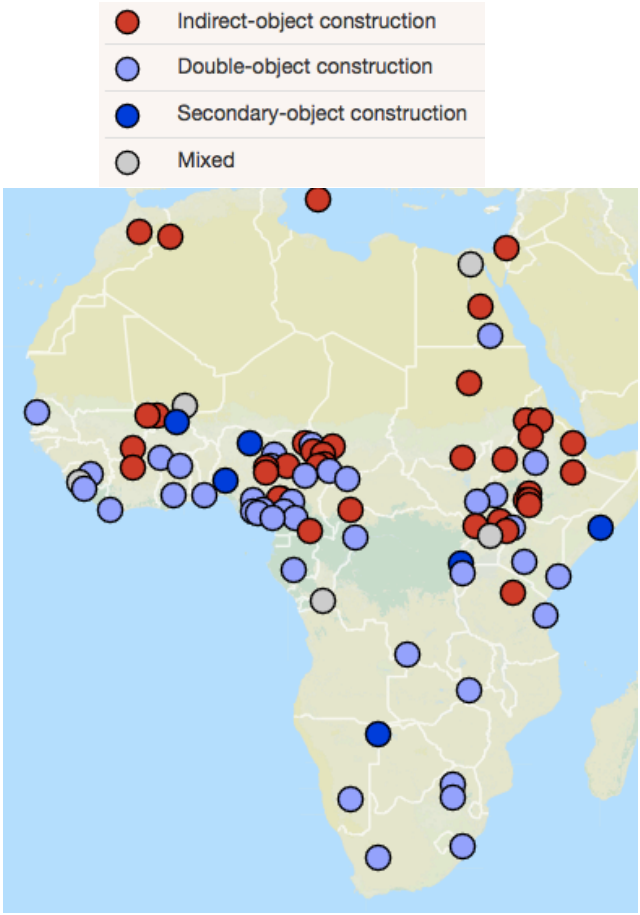
9 A similar approch was adopted by Dryer (1989, 1992) for world-wide samples in the study of language universals. Dryer suggests the unit *genus* which is a level between the individual language and the larger family. A typical example of genera are the subfamilies of Indo-European, e.g. Germanic, Slavic, Celtic, Romance. Dryer's unit *genus* represents a similar level of granularity between the individual language and the larger family.

feature, for instance, all 10 creoles from the English/Macro-Sudan bi-clan show the double-object construction (only concentrating on exclusive double-object constructions represented by single-coloured blue dots on Map 2)[10]. However, no creole from the English/Macro-Sudan bi-clan shows exclusive indirect-object constructions. Such an overall distribution of creoles from the same bi-clan is what we expect from genealogically/areally related languages, namely that they share many linguistic features. The 10 instances of uniform marking (i.e. double-object construction) should therefore not be given the same weight as other creoles with no close relatives. Thus, in the bi-clan distribution, a bi-clan is counted only once if its members show uniform behavior. The English/Macro-Sudan creoles in Table 1 thus contribute 10 points to the language count, but only one point to the bi-clan count. Of the five languages that belong to the French/Macro-Sudan bi-clan, four have double-object constructions, whereas one creole (Louisiana Creole, and here a special geographical variety of Pointe Coupee, Neumann-Holzschuh & Klingler 2013) features the indirect-object construction. Therefore, this bi-clan is counted twice, once for the existence of the double-object construction and once for existence of the indirect-object construction. In this way, we capture the linguistic diversity within and across bi-clans. The advantage of this method is straightforward: Bi-clans are treated alike independently of their size – bi-clans with few creoles have the same impact as bi-clans with many more creoles. The crucial criterion is whether the creoles of one and the same bi-clan show the same or different feature values (for a more detailed discussion of the bi-clan approach, see Michaelis 2016).

If we now apply the bi-clan distribution to all *APiCS* creoles as displayed in Table 1, the majority of double-object marking creoles shrinks and we have a nearly equal split between creoles with exclusive double-object constructions (56%) and those with exclusive indirect-object constructions (44%). Here the bi-clan subdivision helps us to realize that the indirect-object construction is not a minor pattern within creoles world-wide, but also constitutes a widely represented construction type of the world's creoles.

[10] The other 7 creoles from the English/Macro-Sudan bi-clan also predominantly feature the double-object constructions with indirect-object constructions as a minor option (compare the small red slices in the pie-charts on Map 2).

If one considers the geographical distribution, there is a clear-cut division between on the one hand Atlantic and French Indian Ocean creoles, and on the other hand South Asian, Southeast Asian and Pacific creoles. This division does not correlate with lexifier languages as, for instance, English/Macro-Sudan creoles show double-object constructions whereas English/Oceanic creoles (Tok Pisin and Bislama) show indirect-object constructions. The same is true for other European lexifiers. Portuguese/Macro-Sudan creoles (Santome, Principense, Angolar) show double-object constructions whereas Portuguese/Indic (Diu-Indo Portuguese, Korlai, Sri Lanka Portuguese) and Portuguese/Malay creoles (Papiá Kristang, Batavia Creole) show indirect-object constructions. Once we compare the creole data to the corresponding *WALS* data, we clearly see that it is the substrates/adstrates which determine the creole patterns. I will first look at the languages of Africa.



Map 3: Ditransitive Constructions: The Verb 'Give' (Haspelmath 2005)

We see that in sub-Saharan Africa, the clear majority of languages show the double-object construction, as examples (5-6) illustrate.

(5) Wolof (Atlantic; Creissels 2005: 63)

Dama-y	*jox*	*ganaar gi*		*dugub ji*	
VFOC.SM1SG-TAM	give	hen	DEF	millet	DEF
		REC		THEME	

'I'm giving the millet to the hen.'

(6) Fongbe (Kwa; Lefebvre & Brousseau 2002: 254)

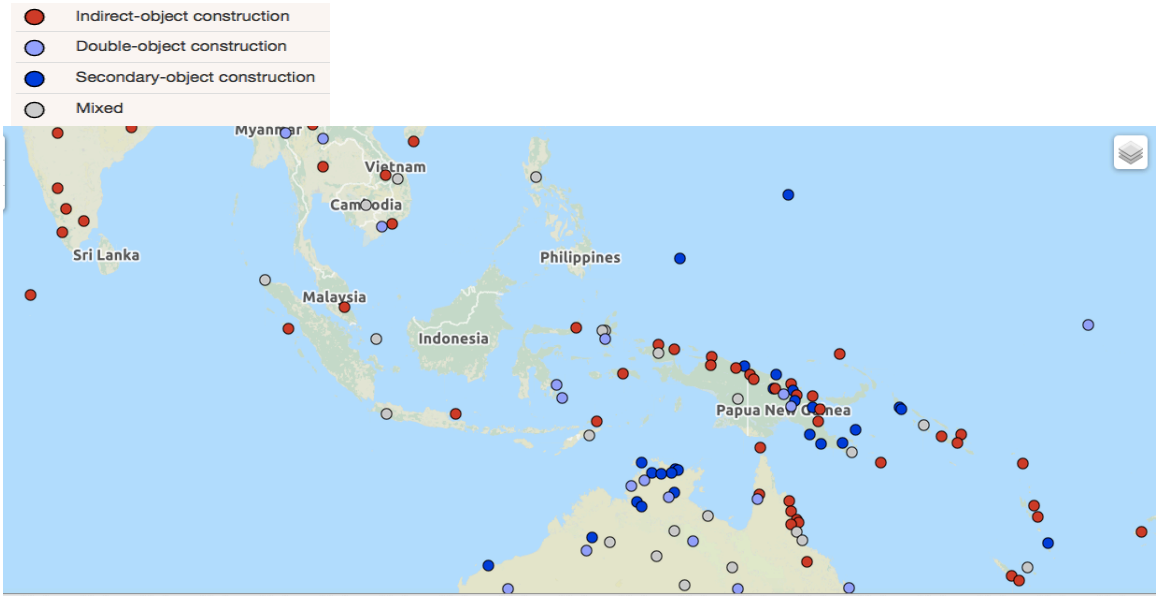
Kɔkú	*ná*	*Asíbà*	*àsɔ́n*
Koku	give	Asiba	crab
		REC	THEME

'Koku gave Asiba (some) crab).'[11]

Most of the languages on Map 3 are from western Africa and central-eastern Africa, because in these areas the diversity is far greater than in the south. Almost all of central-southern Africa is occupied by closely related Bantu languages, and these all show double-object constructions, so adding new dots would not add much new information. We do see three areas where the indirect-object construction is predominant: in Ethiopia, Semitic and Cushitic languages tend to show indirect-object constructions, and in Mali, the Mande languages tend to show indirect-object constructions. But these are all interior regions – note that the coastal regions have exclusively double-object languages. Indeed, these are precisely the regions where most of the slaves which were deported to the various islands close to West and East Africa and the Caribbean came from.

If we now consider the substrate/adstrate languages from South Asia, insular Southeast Asia and the Pacific which are relevant for the corresponding *APiCS* creoles, it is the indirect-object construction which prevails in these regions.

[11] Lefebvre & Brousseau (2002: 254) note that the linearization THEME RECIPIENT is equally possible.



Map 4: Ditransitive Constructions: The Verb 'Give' (Haspelmath 2005)

(7) Kwaio (Oceanic; Keesing 1985: 30)

'E-meru	*meru*	*kwate-a*	*boo*	*ba'ita*	***fa**-na.*
we	1PL.SUBJ	give-3SG.OBJ	pig	big	for-3SG
			THEME		REC

'We gave him a pig.'

(8) Riau Indonesian (Malayic; Gil, p.c.)

Kenapa	*David*	*tak*	*kasi*	*ikan*	***sama***	*dia.*
why	David	neg	give	fish	with	him/her
				THEME	REC	

'Why didn't you (David) give her the fish?'

As the Romance lexifiers show the indirect-object construction, too, one may want to argue for convergence or reinforcement of the lexifer's and the substrate/adstrate's pattern (e.g. for Portuguese-based Korlai and French-based Tayo). But English-based Tok Pisin (see example 4) is striking in this perspective because here only the indirect-object construction is present – reflecting its Oceanic substrates against indirect-object and double-object constructions in its lexifier English – whereas no Atlantic English-based creole has only the indirect-object construction.

Considering these data, I would like to claim (with Michaelis & Haspelmath 2003) that in ditransitive constructions, creoles clearly reflect their substrate/adstrate pattern against possibly conflicting patterns in their lexifiers.

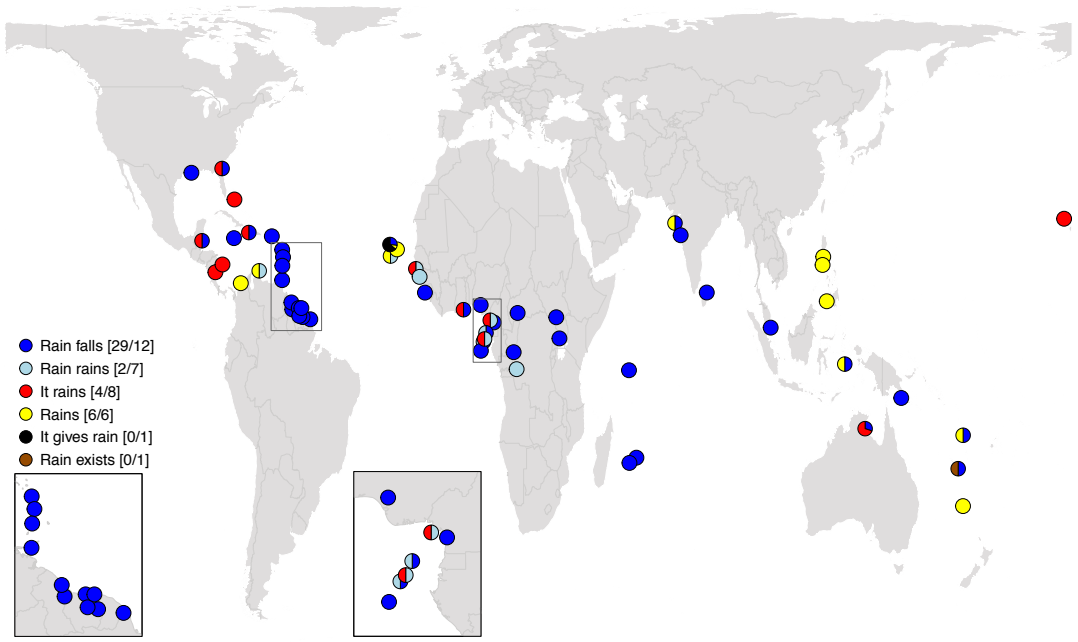
Of course, the strongest evidence for substrates comes from features that are not only absent in the creoles's lexifiers, but also rare in the world's languages. The double-object construction is in fact very common in the world's languages, but the indirect-object construction is about equally common (as can be seen from Haspelmath 2005). However, what makes the present case (and the other constructions still to be discussed) relatively strong is the fact that the explanation works for a *large number of creoles*, and that the cross-linguistic picture of ditransitive constructions is unusually rich due to the typological data available in *WALS*.

4. Weather constructions: 'raining'

The second construction type under investigation concerns a specific meteorological event, namely RAINING. Meteorological events are not expressed uniformly across languages (cf. Eriksen et al. 2010, 2012 for recent cross-linguistic work), and weather constructions also exhibit interesting variation in the creole languages in *APiCS*. Since languages often code different meteorological events with different syntactic constructions ('the sun is shining', 'it is raining', 'there is a thunderstorm', etc.), I consider exclusively the situation 'it is raining'[12]. English and French show constructions with an expletive subject pronoun, as in *it is raining, il pleut*. However, by far the most common type in the creole languages in *APiCS* is represented by a construction where raining is expressed by a word referring to the natural element 'rain' or 'water' in subject position accompanied by a general verb such as 'fall' or 'hit': 'rain falls' (see value 1, blue dots on Map 5) as illustrated in examples (9)-(12).

Some creoles have more than one raining construction represented as slices of a pie-chart according to their relative importance, as Map 5 illustrates.

[12] See the corresponding *APiCS* chapter Michaelis & *APiCS* Consortium 2013a.



Map 5: Raining constructions in 59 creoles of *APiCS* (Michaelis & *APiCS* Consortium 2013a)

(9) Haitian Creole (French-based, Caribbean; Fattier 2013)
Lapli a pral tonbe talè.
Rain DEF FUT.go fall soon
'It will rain very soon.' (lit. 'Rain will fall soon.')

(10) Early Sranan (English-based, Caribbean; van den Berg & Bruyn 2013)
Areen fadomm.
rain fall
'It is raining.' (lit. 'Rain is falling.')

(11) Mauritian Creole (French-based, Indian Ocean; Baker & Kriegel 2013)
lapli pe tombe
rain PROG fall
'It is raining.' (lit. 'Rain is falling.')

(12) Tayo (French-based, Pacific; Ehrhart & Revis 2013)
lapli le to:mbe
rain SI fall
'It is raining.' (lit. 'Rain is falling.')

It seems as if the 'rain falls' construction is not areally restricted as we find it spread over many different regions of the world with creoles based on various lexifier languages. But as we will see in the more detailed discussion of this feature below, it is very likely that we can detect substrate influence in this feature, too. Unfortunately, raining constructions are not covered by *WALS*, and the typological data available (Eriksen et al. 2010, 2012) can only hint at areal distributions of the various possible constructions.

A very similar construction type to the one in (9-12) is illustrated by example (13) where a 'rain' noun in subject position is combined with a 'rain' verb, i.e. a verb that exclusively (apart from metaphorical usages) refers to raining situations. Only few creole languages show this construction.

(13) Guinea-Bissau Creole (Portuguese-based; Intumbo et al. 2013)
Tcuba na tcubi.
Rain PROG rain
'It is raining.'

Here, the geographical restriction to Africa (and some nearby islands) is striking. Besides Kikongo-Kituba and Guinea-Bissau Kriyol, which have only this 'raining' construction, it is found as one option in Cape Verdean Creole of Brava, Casamancese Creole, Pichi, the three Gulf of Guinea creoles Principense, Santome, and Angolar, and Papiamentu, the only Caribbean creole with this construction outside of Africa.

The third construction tpye ('it rains') consist of an expletive subject and a 'rain' verb, as is well documented in Western European languages like English *(it rains*), French *(il pleut*), and Dutch (*het regent*).
(14) Hawai'i Creole (English-based; Velupillai 2013)
its *gɔna ɹeɪn*
3SG.is FUT rain
'It's going to rain.'

Interestingly, most of the 12 languages with this value are English-based languages, and all four languages that have only this construction are also English-based languages (Bahamian Creole, Nicaraguan Creole English, San Andres Creole English, and Hawai'i Creole). These creoles can be classified as more acrolectal and as such more closely reflecting the English expletive construction. But some Portuguese-based languages too (Santome, Guinea-Bissau Kriyol, Casamancese Creole) have the construction with an expletive subject, even though their lexifier Portuguese does not show an expletive subject in the corresponding construction (*chove* 'it rains'):

(15) Santome (Hagemeijer 2013)
Ê *ka* *sôbê* *muntu fan*.
3SG IPFV rain a.lot PCL
'It rains a lot!'

In the last construction type, there is just a 'rain' verb without any subject (as in Spanish *llueve* and Portuguese *chove* 'it rains').

(16) Ternate Chabacano (Spanish-based; Sippola 2013)
Ta *yobé*.
IPFV rain.V
'It rains'. OR: 'It's raining.'

(17) Ambon Malay (Austronesian; Paauw 2013)
Mo ujang.
FUT rain
'It's going to rain.'

Here, we again see some geographical patterning: The 12 creole languages with this value are spoken on the Cape Verde Islands, in Asia, and the Pacific, but not in mainland Africa, the Gulf of Guinea, the Caribbean (except for Papiamentu and Palenquero), or Australia.

Table 2: Raining constructions in creoles (exclusive marking only)

	APiCS creoles		*APiCS* bi-clans of creoles	
rain falls	29	70%	13	**61%**
rain rains	2	5%	2	**10%**
it rains	4	10%	2	**10%**
rains	6	15%	4	**19%**
total	41	100%	21	100%

Table 2 illustrates the distribution of the different raining constructions over the creoles in *APiCS*. Again, for simplicity reasons, we restrict ourselves to creoles with exclusive feature values, i.e. which only show one construction in this feature. As we have seen before, the first construction 'rain falls' is by far the most widespread in creoles, and this holds in the bi-clan distribution, too. Out of the 16 creoles belonging to the English/Macro-Sudan bi-clan, for instance, 14 feature the 'rain falls' construction[13]. As there is no *WALS* chapter relating to raining construction, we rely on Eriksen et al. (2010, 2012), who investigate weather constructions in some 30 languages of different world regions. They observe that the 'rain falls' construction is also widespread world-wide. Despite this fact, I think that there is good evidence for substrate influence in this particular 'raining' construction. Koopman (1986) and Lefebvre (1998) have pointed out the parallel syntactic constructions in West African languages, in particular in Fongbe, which is one of the most important substrates for Haitian Creole (cf. example 9) and other Macro-Sudan-clan creoles:

(18) Fongbe (Kwa; Lefebvre & Brousseau 2002: 245)
Jí jà
rain fall
'It is raining.' (lit. 'Rain is falling.')

Similar constructions are found in Bantu languages, giving rise to the same structures, for

[13] In some of these languages, the 'rain falls' construction is shared with the expletive construction 'it rains'.

instance, in the French-based creoles of the Indian Ocean (see examples 11 from Mauritian Creole).

(19) Swahili (Bantu; Vitale 1981: 57f.)
(*Mvua) i-na-nyesha*.
(rain) 9-PRS-fall
'It is raining.' (lit. 'Rain is falling.')

For the Australian English-based creole language Kriol, Schultze-Berndt & Angelo (2013) also argue for substrate influence in the same construction type (*rein bin-a bol-dan*, [rain PST-POT fall-down] 'lit. 'rain falls'), as a similar construction is found in at least some of the traditional languages of the area (see also Kofod (1978:195) on Miriwung, East Kimberley):

(20) Jaminjung (Mirndi, Northern Australia; Schultze-Berndt 2013)
Gugu gardambiyang wirlarrungburru.
gugu ga-rda-m=biyang wirlarrung-burru
water 3SG.S-fall-PRS=SEQ lightning-PROPR
'It is raining now (lit. 'water/rain falls'), with lightning.'

The same picture holds for Tok Pisin, Bislama, and Tayo which have Oceanic substrates which equally show 'rain falls' constructions:

(21) Mwotlap (Oceanic, Northern Vanuatu; François 2001: 342)
Na-smal me-smal.
ART-rain PFV-rain
'It is raining.' (lit. 'Rain is raining')

(22) Xârâcùù (Oceanic, New Caledonia; Moyse-Faurie 2013)
kwiè xwa
rain fall
'It rains.' (lit. 'Rain falls.')

But it is not only for the 'rain falls' type that I would like to postulate substrate influence. The 'rains' type (yellow dots) which is found, for instance, in all Chabacano varieties and in Ambon Malay is also widespread amongst languages of insular Southeast Asia and the Philippines, and – as it seems – in some areas of mainland Southeast Asia which are substrate/adstrate languages to the cited creoles.

(23) Tagalog (Malayo-Polynesian, Northern Philippines; Kroeger 1993: 49)
umuulan
IPFV.rain
'It is raining.'

(24) Jakarta Indonesian (Malayo-Polynesian; Conners & Gil 2013)
Ujan.
rain
'It's raining.'

(25) Mandarin Chinese (Sino-Tibetan; Zhang 2013)
下雨了。
Xià yǔ le.
Xiayu le
rain PRS
'It is raining.'

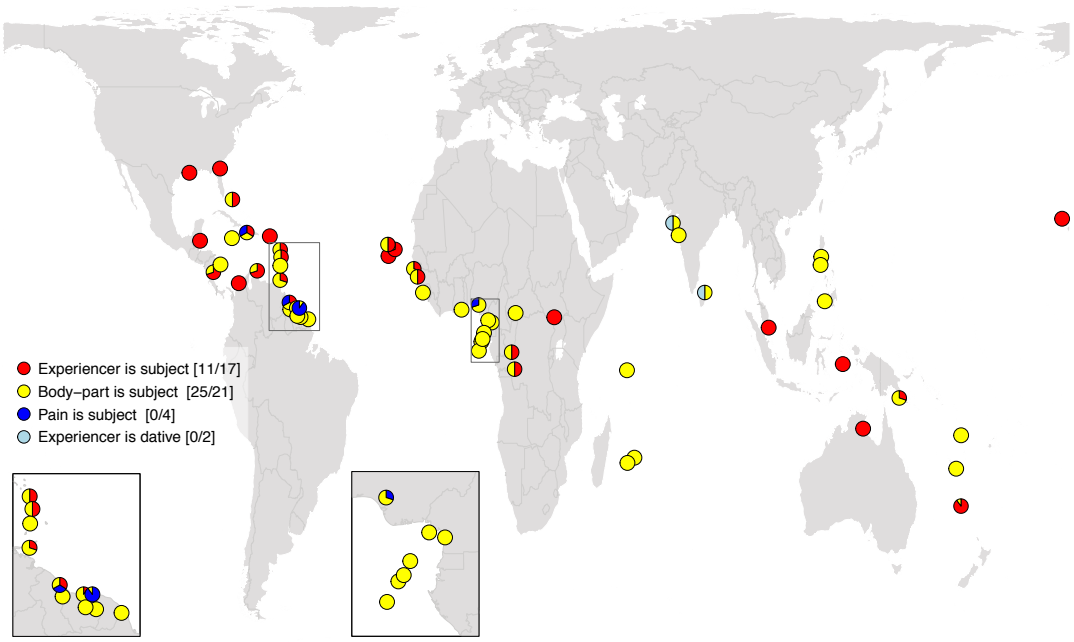
It is interesting that besides the cited creole languages, we find this construction also in other *APiCS* languages of the same area, such as the two English-based languages Singlish (besides the expletive construction most probably inherited from English) and Chinese Pidgin English, and Singapore Bazaar Malay. This additional data makes the substrate claim even stronger.

As an interim conclusion, we can state that in "raining" constructions, too, the valency patterns of creoles (overwhelmingly) go back to their substrates/adstrates, and not to their lexifiers. The fact that in this construction type not only single creoles, but creole languages world-wide mirror their substrate pattern again strengthens my claim.

5. Experiencer constructions 'headache'

In this section, I will look at a third construction type in creole languages, namely EXPERIENCER CONSTRUCTIONS. As experiencer constructions vary significantly depending on the more special type of experience they express (e.g. pain, emotions, fear), I will concentrate on pain sensations and more specifically on situations which involve 'headache', as in English *She has a headache* (cf. the corresponding *APiCS* chapter Michaelis & *APiCS* Consortium 2013b). Such headache constructions involve three participants, an experiencer, who experiences the pain sensation ('she'), the sensation itself ('ache'), and potentially a body-part ('head') to which this sensation is related. There is a lot of interesting variation in such constructions. However, here I focus on the question which of the three participants is coded in subject position. Subject is defined here as an argument that is coded like the typical agent in a monotransitive clause, or the single argument of an intransitive clause.

Map 6 shows the distribution of the different types of 'headache' constructions in the 59 creoles in *APiCS*. If a language has more than one feature value, the different construction types are rendered as slices of a pie-chart according to their relative importance, as already shown in earlier maps. As one can see in Map 6, quite a few creoles show more than one construction type.



Map 6: Experiencer constructions 'headache' in 59 creoles of *APiCS* (Michaelis & *APiCS* Consortium 2013b)

In the first construction type, which is well-known from the European lexifiers ('She has a headache', red dots on Map 6) the experiencer is coded as subject, as in examples (26) and (27). This type is fairly widespread among the *APiCS* languages.

(26) Kriol (English-based; Schultze-Berndt & Angelo 2013)
Maitbi ***im*** *hedeik* (...)
maybe 3SG headache [...]
'Maybe he has got a headache (...)'.

(27) Papiá Kristang (Portuguese-based; Baxter 2013)
Yo *teng* *due* *kabesa*
1SG have hurt head
'I have a headache.'

In the second type, the body-part is coded as the subject of the sentence, as in 'Her head is aching' (yellow dots on Map 6). This type seems to be the most prominent construction in the *APiCS* sample. Within the creole languages showing this feature

value, there are several subtypes. The experiencer can be retrievable via the oblique object of the verb 'eat', as in example (28) from Saramaccan, but it can also be expressed as a possessor of the body-part noun, as shown in example (29) from Angolar.

(28) Saramaccan (English-based; Aboh et al. 2013)
Edi ta njan a mi.
head ASP eat PREP 1SG
'I have a headache.' (lit. 'The head is eating me.')

(29) Angolar (Portuguese-based; Maurer 2013a)
N'tê m tha ruê.
head my PROG ache
'I have a headache.' (lit. 'My head is aching.')

Ghanaian Pidgin English shows a third subtype: the experiencer is expressed both through the possessive pronoun *ma* and the object pronoun *mi* .

(30) Ghanaian Pidgin English (English-based; Huber 2013)
mà hɛd dè pen/nak mi
1SG.POSS head PROG pain/knock 1SG.OBJ
'I have a headache.' (lit. 'My head is hurting/ knocking me.')

In yet another variant of the construction type 'body-part is subject', the experiencer is not expressed at all and has to be inferred from the linguistic or extra-linguistic context. One example comes from Diu Indo-Portuguese.

(31) Diu Indo-Portuguese (Portuguese-based; Cardoso 2013)
Kabes tə dw-e.
head IPFV.NPST hurt-INF
'(My) head is hurting.'

The construction type corresponding to the next value is only marginally represented within the *APiCS* sample (dark blue dots on Map 6) and it exists only as an alternative

option besides another construction. Here the 'pain' itself is coded as subject, as in 'Headache is affecting her'. An example comes from Sranan:

(32) Sranan (English-based; Winford & Plag 2013)
Ede-hati *e* *kiri* *mi.*
head-hurt IPFV kill me
(lit.) 'A headache is killing me.'

In the last construction type (light blue dots on Map 6), the experiencer is marked as dative, that is, like the recipient of a typical ditransitive verb like 'give'. In the following example (33), the preposition *a*, which otherwise occurs on recipients, marks the human experiencer.

(33) Diu Indo-Portuguese (Portuguese-based; Cardoso 2013)
A ***mi*** *tə* *siti* *dor* *də* *kabes.*
DAT 1SG.OBL IPFV.NPST feel.INF pain of head
'I have a headache.' (lit. 'To me there is feeling a headache.')

The 'experiencer is dative' construction occurs only marginally, too, but has a clear South Asian areal pattern (especially if one considers also other experiencer constructions as studied in *APiCS*, e.g. experiencer constructions with 'like' and 'fear', see Michaelis & *APiCS* Consortium 2013c, 2013d).

Table 3 summarizes the figures for the experiencer constructions. To keep the picture somewhat simple, again only languages with exclusive marking are represented. The numbers for 'pain is subject' and 'experiencer is subject' constructions thus do not show up, as they are shared feature values and occur only as options besides other construction types (for a full picture see Michaelis & *APiCS* Consortium 2013b).

Table 3: Experiencer constructions with 'headache' (exclusive marking only)

	APiCS creoles		*APiCS* bi-clans of creoles	
Experiencer is subject	11	31%	10	**48%**
Body-part is subject	25	69%	11	**52%**
Total	36	100%	21	100%

Note: 'Pain is subject' and 'Experiencer is dative' are not represented here because they are only shared options besides other construction types.

Considering the figures in Table 3 for the *APiCS* creoles, one might be led to think that the 'body-part' construction is much more prominent across creoles than the 'experiencer is subject' construction as it is present in 69% of all creoles in *APiCS*. But in the bi-clan distribution, uniformly coded languages of the same bi-clan are counted only once (for instance, 7 Macro-Sudan/English creoles show the 'body-part' construction and count only once), and therefore there is a nearly equal split between the two construction types: 52% 'body-part is subject' vs. 48% 'experiencer is subject'. Thus both constructions are well represented among creole languages.

What is the origin of the different construction types in the creoles? It is interesting to note that many *APiCS* creoles with English, Dutch, and French lexifiers only have body-part subjects (for instance, Ghanaian Pidgin English, Angolar, Nengee, Jamaican, Seychelles Creole, Tayo). This option is at best one possible construction in these European lexifiers, but it is certainly not the most prominent construction, as the most neutral way of referring to this experience is to express the experiencer in subject position, as in *I have a headache*, French *J'ai mal à la tête* [I.have pain at the head]. This leads me to suspect substrate influence also in this feature. Unfortunately, there is no *WALS* chapter on experiencer constructions, and the few systematic typological studies available focus on European languages (see Bossong 1998, Haspelmath 2001). For creoles with West African substrates, Ameka's (1990) description of experiencer constructions in Ewe (a Kwa language of West Africa) is of great interest. Ameka (1990: 165) shows that the body-part subject in 'headache' situations is the only construction type available in this language.

(34) Ewe (Kwa; Ameka 1990: 165)
ta vé kofí
head pain Kofi
'Kofi had a headache.' (lit. 'head pained Kofi')

Lefebvre (1998: 251) cites a structurally very similar example from Fongbe, a closely related language, where the body-part is subject, with the experiencer encoded as the object of the verb 'eat'.

(35) Fongbe (Lefebvre 1998: 251)
Tà ɖǔ mi.
head eat me
'I have a headache.' (lit. 'head is eating me.')

In Swahili, a Bantu language, such experiences are also expressed by a construction where the body-part is subject:

(36) Swahili (Madan 1902, s.v.)
kichwa ch-a-ni-uma
head 3SG.SBJ-PRS-1OBJ.hurt
'I have a headache.' (lit. 'head is aching me.')

Thus, quite a few of the Atlantic and French-based Indian Ocean creoles showing the body-part pattern are most probably influenced by their substrates.

There is further good evidence that another construction type was calqued on a substrate/adstrate pattern: the 'experiencer is dative' construction, which is geographically restricted. The two creole languages with this pattern are both located in South Asia: Diu Indo-Portuguese (see example 33) and Sri Lanka Portuguese[14]. If one looks at the relevant substrate/adstrate languages, one finds that the corresponding pattern of dative-marked experiencers is widespread in South Asian languages (see examples 37, 38). At the same time such a construction is not very common in the world's languages, and

[14] A third language in *APiCS* with this feature value is the mixed language Sri Lankan Malay (cf. Slomanson 2013).

moreover not present in the corresponding lexifier language Portuguese, where the experiencer is coded as subject (cf. *Tenh-o dor de cabeça* [have.1SG pain of head]). In the literature this construction is often called "dative subject construction" (see also Michaelis & *APiCS* Consortium 2013c on experiencer constructions with 'like').

(37) Gujarati (Indo-Aryan; Cardona 1965: 110)
məhne *daNtmaN* *dukhe* *che*
1SG.DAT tooth be.painful AUX
'I have a toothache.' (lit. 'To me the tooth is painful.')

(38) Malayalam (Nizar 2010)
eni-kkə *talaveedana* *uɳtə*
1SG-DAT headache be.PRS
'I have a headache.' (lit. 'To me is a headache.')

The marking of experiencers in the South Asian *APiCS* languages by this kind of dative case or adposition can thus clearly be traced back to their substrates/adstrates.

What about the 'experiencer is subject' construction which is well represented in the creoles in *APiCS*? Should we attribute all cases of this feature value in creoles to their (Western European) lexifiers? It may well be that in some of the more acrolectal French-/English-based creoles with this construction, such as Louisiana Creole or Gullah, this interpretation is valid.

But for languages such as Kriol and Papiá Kristang (see examples 26-27), the situation is different. One of Kriol's adstrate languages is Jaminjung, a language of Northern Australia. In this language, it is also the experiencer which is coded as subject in constructions with headache (Schultze-Berndt, p.c.). So the lexifier English (*I have a headache*) and Jaminjung show the same pattern. The situation is similar for Papiá Kristang, whose most important substrate/adstrate language is colloquial Malay of Malacca. According to David Gil (p.c.), the headache experience can be expressed as in (39) in colloquial Malay:

(39) Colloquial Malay (Gil, p.c.)

Aku	*ada*	*sakit*	*kepala*
1SG	exist/have	hurt	head

'I have a headache.'

The predicate *ada* is used in existential constructions, but also in predicative possessive constructions. So it may well be that this colloquial Malay construction is the source of the 'experiencer is subject' construction in Papiá Kristang (see example 27). It so happens that Portuguese has a very similar construction with *Tenho dor de cabeça* 'I have pain of head'.

Generalizing over cases like Kriol and Papiá Kristang, it becomes clear that the creole patterns ultimately mirror substrate/adstrate patterns which themselves may happen to coincide with the lexifier ones (see the same line of argumentation for cases of indirect-object marking cited in §3). The languages showing 'experiencer is subject' pattern are therefore consistent with the present claim that in 'headache' constructions, too, it is the substrate/adstrate pattern which overwhelmingly prevails in creoles world-wide.

6. Motion-to and motion-from constructions

The last construction type to be discussed are motion-to and motion-from constructions. Here the question is whether creole languages use the same strategy or different strategies to express the two opposite orientations: (i) motion to a reference object (GOAL, 'I go to the market'), and (ii) motion from a reference object (SOURCE, 'I come from the market'). The reference objects in these constructions are frequent places like 'home', 'town', 'village', 'market', or 'forest' (cf. the corresponding *APiCS* chapter Michaelis & *APiCS* Consortium 2013e).

All Western European lexifier languages have different constructions for motion-to and motion-from, using different prepositions: English *to town/from town*, French *à la maison/ de la maison*, Portuguese *ao mercado/do mercado*. But in *APiCS*, it is striking to see that many European-based creoles do not follow the Western European lexifier

patterns and instead mark goal and source identically, as illustrated in examples (40a-b) from Seychelles Creole and (41a-b) from Krio.[15]

(40) Seychelles Creole (French-based; Michaelis & Rosalie 2013)

a. motion-to: *mon* *al* ***dan*** *bwa*
1SG go in forest
'I go into the forest.'

b. motion-from: *mon* *sorti* ***dan*** *bwa*
1SG come.from in forest
'I come out of the forest.'

(41) Krio (English-based; Finney 2013)

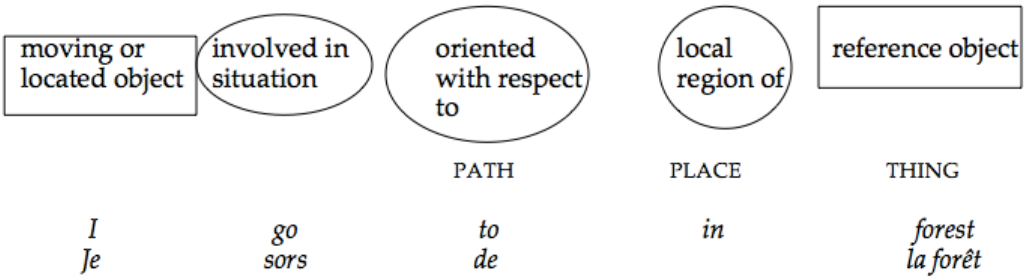
a. motion-to: *a* *di* *go* ***na*** *di* *makit*
1SG PROG go LOC ART market
'I am going to the market.'

b. motion-from: *A* *jɛs* *kɔmɔt* ***na*** *di* *makit*
1SG just come LOC ART market
'I just came back from the market.'

It is not the case that the creoles cited do not use any prepositions, but the interesting fact is that they use the *same* preposition to refer to both opposite orientations. Appartently, the prepositions refer to something else than the prepositions *to, from, à* and *de* in the lexifier languages. To better understand the differences between Seychelles Creole and its lexifier French, and Krio and its lexifier English, I would like to refer to Figure 2, where a local situation is decomposed.

[15] Throughout this section, the (a) examples show motion-to constructions and the (b) examples show motion-from constructions.

Figure 2. Structure of a local situation (Jackendoff 1983: 161ff.; Lehmann 1992: 629)

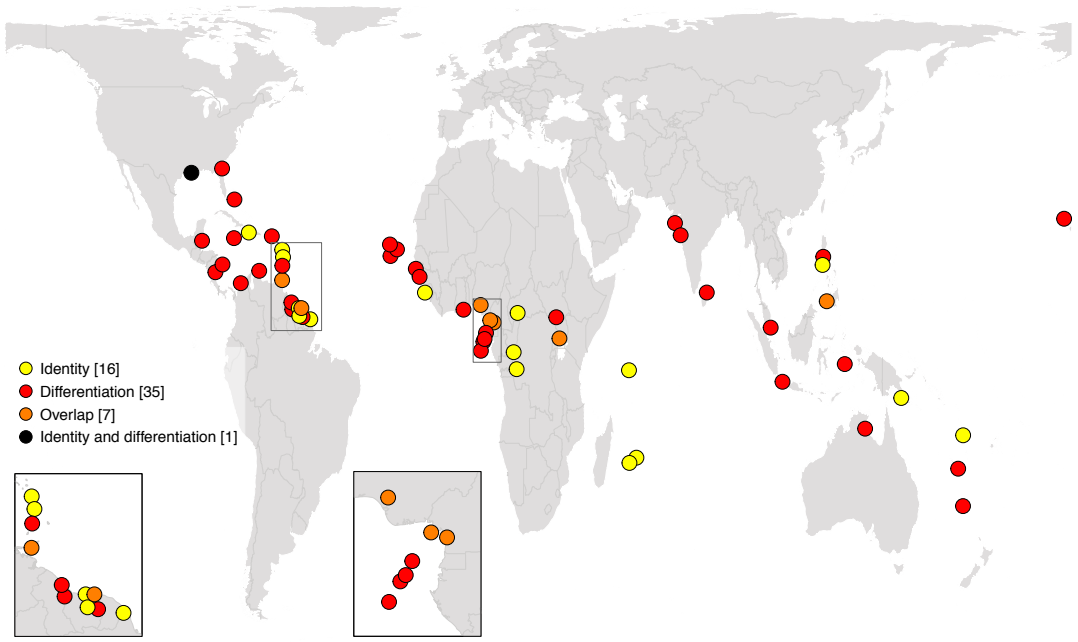


A moving or located object ('I') is involved in a situation ('go') which is oriented ('to') with respect to a local region ('in') of a reference object ('forest'). In English, the two semantic components 'oriented with respect to' (PATH) and 'local region of' (PLACE) are fused into one preposition *into* (which shows a somewhat unexpected linearization, first PLACE *in-* and then PATH *-to*). Within the PATH segment, the motion-to and motion-from concepts are marked differently, *into* vs. *out of*. The same is true for French: motion-to is marked by the preposition *à*, whereas motion-from is marked by the preposition *de*.

(42)	English	French
a. motion-to	*I go **into** the forest.*	*Je vais **dans** la forêt.*
b. motion-from	*I come **from/out of** the forest.*	*Je sors **de** la forêt.*

If one now compares the corresponding examples from Seychelles Creole and Krio to the English and French ones, it becomes obvious that in the creole languages the prepositions *dan* and *na* do not refer to path, but to place or local region of the reference object, the 'forest' or the 'market' in our examples. Thus, these prepositions are the same in both orientations as they are not involved in the expression of path (see also Wälchli & Zuñiga 2006: 292 for Haitian Creole, where we find similar examples). It is from the meaning of the verbs *go* and *come* that the hearer has to infer the relevant orientation.

As illustrated in Map 7, 16 creoles code motion-to and motion-from constructions identically (yellow dots).



Map 7: Motion-to/motion-from in 59 creoles of *APiCS* (Michaelis & *APiCS* Consortium 2013e)

35 creoles in *APiCS* code motion-to and motion-from differently (red dots on Map 7). This can be achieved through three coding subtypes. First, a language has two different prepositions to mark motion-to and motion-from, as also seen in the European lexifiers cited above. This is the case, for instance, in Bahamian Creole, Palenquero, and Batavia Creole.

(43) Batavia Creole (Portuguese-based; Maurer 2013b)

a. *Ile anda* ***nu*** *chang* [...].
he go to land
'He went to the property [...].'s

b. ***Di*** *undi sta bi?* — ***Di*** *otër tera.*
from where IPFV come – from other country
'Where do you come from? — From abroad.'

The second subtype of this feature value consists of an OPTIONAL adposition to express motion-to, but a different obligatory adposition to express motion-from. Here we find

languages like Creolese, Papiamentu, two Cape Verdean Creole varieties, Kriol, and Sri Lanka Portuguese.

In a third subtype of the value "differentiation", motion-to is never marked, whereas motion-from is always marked by an adposition or a serial verb. Jamaican, Casamancese Creole, Guinea-Bissau Kriyol, Papiá Kristang, and Tayo have prepositions (*frahn, di, de*) for expressing motion-from (see example 44). Sri Lanka Portuguese shows postpositions, and the Gulf of Guinea creoles (Angolar, Principense, Santome, and Fa d'Ambô) mark motion-from with the serial verb *fõ/fo* 'come from' (see example 45).

(44) Casamancese Creole (Portuguese-based; Biagui & Quint 2013)

a. *Mariya bay fera.*
   Mary go market
   'Mary went to the market.'

b. *Mariya beŋ* ***di*** *fera.*
   Mary come from market
   'Mary came back from the market.'

(45) Principense (Portuguese-based; Maurer 2013c)

a. *N we fya.*
   1SG go market
   'I went to the market.'

b. *N vika* ***fo*** *fya.*
   1SG come come.from market
   'I came from the market.'

In seven languages there is overlap in the coding of motion-to and motion-from constructions (orange dots on Map 7). In Cameroon Pidgin English (Schröder 2013), for instance, the preposition *fo* 'in, at, on, to, from' can be used in both orientation contexts, whereas motion-to can additionally be expressed without any marker.

Table 4: Motion-to and motion-from in creoles (exclusive marking only)[16]

	APiCS creoles		*APiCS* bi-clans of creoles	
Identity	16	28%	7	**28%**
Differentiation	35	60%	14	**56%**
Overlap	7	12%	4	**16%**
Total	58	100%	25	100%

As Table 4 illustrates, 16 creole languages mark motion-to and motion-from identically. On top of that, there are 7 more creole languages that show overlap, i.e. they may have identical marking of motion-to and motion-from (besides differentiation, as in Cameroon Pidgin English, mentioned above). I subsume these languages under the group of identity marking because, given that all Western European lexifiers show the differentiation pattern, the possibility to mark both orientations identically is remarkable and is in need of being explained. If we consider the bi-clan distribution and add the bi-clans showing overlap (16%) to the ones showing identity (28%), we end up with a total of 44% bi-clans that may show the identity pattern.

Interestingly, we observe a split between English- and French-based Atlantic/Indian Ocean creoles: All French-based creoles show the identity pattern, whereas the English-based Atlantic creoles mostly show differentiation. It is only the English-based Surinamese creoles ((Early) Sranan and Saramaccan), Trinidad English Creole, and the West African English-based languages which show identical coding or overlap. Other Caribbean English-based creoles, such as Jamaican, Creolese, Belizean Creole, and Gullah, follow the English differentiation pattern.

How can we explain the identity pattern which is not available in the European lexifier languages? Even though there is no *WALS* map for this feature, we have good cross-linguistic data on intransitive motion-to/motion-from constructions (Wälchli & Zuñiga 2006). First, in many West African and Bantu languages, which are substrates of the Atlantic and Indian Ocean contact languages, motion-to and motion-from in sentences relating to a situation like 'I go to/come from the market' are not overtly marked, but orientation is expressed through the semantics of the verb (see Michaelis

[16] One minor value ("identity and differentiation") has been omitted (for the full picture, see Michaelis & *APiCS* Consortium 2013e).

2008, Wälchli & Zúñiga 2006: 292ff.). If prepositions are used, they do not refer to orientation (motion-to/motion-from), but to the local region of the located object (see Figure 2). This pattern can be found in a great variety of Niger-Congo languages, such as Ewe, Twi, Efik, Bambara, Zulu, Swahili. In Swahili, the postnominal locative case marker *-ni* indicates the local region of the reference object and not path. Thus, *-ni* closely mirrors the prepositions in the creole examples (40-41).

(46) Swahili (New Testament, Mark 1, 11; 19)

a. *a-ka-enda* *bahari-ni*
   3SG-SEQ-go sea-LOC.in
   'He went to the sea'.

b. *sauti* *i-ka-toka* *mbingu-ni*
   voice 3SG-SEQ-come.from heaven-LOC.in
   'A voice came from heaven'.

Second, in Philippinic languages, e.g. Tagalog, which are substrate/adstrates to the Chabacano varieties, goal and source may not be overtly marked either, as is the case in Ternate Chabacano and Zamboango Chabacano (whereas Cavite Chabacano variety has different markers for goal and source).[17]

However, there are more than 50% of the creole bi-clans that differentiate between motion-to and motion-from. Do these languages follow their Western European lexifier patterns? Here I will argue again that the picture is more complex. If we have a closer look at the relevant contact situations, we see that in the overwhelming number of cases, the relevant substrate/adstrate languages also have different means to express motion-to and motion-from. Tayo, for instance, has Oceanic substrates which differentiate motion-to and motion-from prepositions (Rivierre 1980: 220, 351, Osumi 1995: 80f., Bril 2002: 296, 309). The same holds for English-based Kriol, which has differentiating

[17] It is not clear to me why so many English-based Atlantic creoles with West-African substrates do not show the identity pattern. Interestingly, the English-based Surinamese creoles reflect the substratal identity pattern. It is not only in this feature that we see this different behavior in the English-based Atlantic creoles (for a similar picture see the *APiCS* feature "noun phrase conjunction and comitative", http://apics-online.info/parameters/71#2/13.9/10.2). More research is needed here.

substrates/adstrates (e.g. Nunggubuyu, cf. Wälchli & Zuñiga 2006: 300). One important substrate of Papiá Kristang and Batavia Creole, Malay, has different goal/source marking, too. Korlai, Diu Indo-Portuguese and Sri Lanka Portuguese all have differentiating adstrate languages. In all these cases, it could be argued that the creole languages just continue the pattern of their French, Portuguese or English lexifiers, which differentiate both orientations. But as already noted for previous construction types, I would argue again that these examples are consistent with my claim that it is the substrate/adstrate pattern that prevails over the lexifier pattern. If it happens that both lexifier and substrate have the same typological feature value, no positive evidence can be adduced. However, these cases do not invalidate the overall generalization that also motion-to/motion-from constructions mirror the substrate patterns instead of the lexifier patterns of the creoles.

Conclusion

In this paper, I have shown that in the domain of valency, creoles show a substantial amount of constructional calquing from their substrates. In all four construction types (ditransitive constructions (§3), weather constructions (§4), experiencer constructions (§5), and motion constructions (§6)), I was able to demonstrate recurrent matches between substrate and creole structures (contrasting with lexifier structures) in different regions of the world and with different substrates. Such a rich picture, which became possible with the publication of the *Atlas of Pidgin and Creole Language Structures* (OUP 2013), excludes the possibility of an accident and makes substrate influence the key in explaining the creole data. This shows that during the process of creolization the creators of creoles systematically calqued the valency patterns of their dominant languages (substrates) onto the emerging creoles.

I am very much aware of the fact that the present paper is only the first step in a much more ambitious endeavour, namely to undertake a large-scale meticulous qualitative world-wide comparison of creoles with their contributing languages against the background of the world-wide diversity of languages. Such a comparison will tell us then *which* structural features are prone to be calqued during creolization and which are

not (for instance, creole word order patterns seem to mirror their lexifier patterns more faithfully, cf. Lefebvre 2015; Michaelis & Haspelmath 2014)[18].

References

Aboh, Enoch O. & Veenstra, Tonjes & Smith, Norval S.H. 2013. Saramaccan structure dataset. In Michaelis et al. (2013c).

Alleyne, Mervyn C. 1980. *Comparative Afro-American — An Historical-Comparative Study of English-Based Afro-American Dialects*. Ann Arbor: Karoma.

Ameka, Felix. 1990. The grammatical packiging of experiencers in Ewe: A study in the semantics of syntax. *Australian Journal of Linguistics* 10: 139-81.

Baker, Philip & Kriegel, Sibylle. 2013. Mauritian Creole dataset. In Michaelis et al. (2013c).

Bakker, Peter. 2008. Pidgins versus creoles and pidgincreoles. In Kouwenberg, Silvia & Singler, John Victor (eds.), *Handbook of Pidgin and Creole Studies*, Oxford: Wiley-Blackwell, 130-57.

Bakker, Peter, Aymeric Daval-Markussen, Mikael Parkvall & Ingo Plag. 2011. Creoles are typologically distinct from non-creoles. *Journal of Pidgin and Creole Languages* 26(1). 5–42. doi:10.1075/jpcl.26.1.02bak.

Baxter, Alan. 2013. Papiá Kristang structure dataset. In Michaelis et al. (2013c).

Biagui, Noël Bernard & Nicolas Quint. 2013. Casamancese Creole structure dataset. In: Michaelis et al. (2013a).

Bickerton, Derek. 1981. *Roots of language*. Ann Arbor: Karoma.

Boretzky, Norbert. 1983. Kreolsprachen, Substrate und Sprachwandel, Wiesbaden: Harrassowitz.

Bossong, Georg. 1998. Le marquage de l'expérient dans les langues d'Europe. In Feuillet,

[18] See Muysken & Smith (2015) for a very detailed picture of which of the creole features of English-based Surinamese creoles were calqued from their West African substrate languages Gbe and Kikongo and which were apparently not calqued.

Jack (ed.). *Actance et valence dans les langues d'Europe*. Berlin: Mouton de Gryuter. 259-94.

Bril, Isabelle. 2002. *Le nêlêmwa (Nouvelle-Calédonie): Analyse syntaxique et sémantique*. Paris: Peeters.

Cardona, George. 1965. *A Gujarati Reference Grammar*. Philadelphia: The University of Pennsylvania Press.

Cardoso, Hugo C. 2013. Diu Indo-Portuguese structure dataset. In Michaelis et al. (2013c).

Chaudenson, Robert. 1992. *Des îles, des hommes, des langues*. Paris: L'Harmattan.

Conners, Thomas J. & Gil, David 2013. *Jakarta Indonesian Valency Patterns*. In Hartmann, Iren & Haspelmath, Martin & Taylor, Bradley (eds.) 2013. *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://valpal.info/languages/jakarta-indonesian/examples/19, Accessed on 2015-09-21)

Creissels, Denis. 2005. A typology of subject and object markers. In Voeltz, Erhard F. K. (ed.), *Studies in African linguistic typology*. Amsterdam: Benjamins, 43-70.

Daval-Markussen, Aymeric. 2014. First steps towards a typological profile of creoles. *Acta Linguistica Hafniensia* 45(2):1-22.

Dryer, Matthew & Haspelmath, Martin (eds.). 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info)

Ehrhart, Sabine & Revis, Melanie. 2013. Tayo structure dataset. In: Michaelis et al. (2013a).

Eriksen, Pål & Kittilä, Seppo & Kolehmainen, Leena. 2010. The Linguistics of Weather: Cross-Linguistic Patterns of Meteorological Expressions. *Studies in Language* 34.3:565–601.

Eriksen, Pål & Kittilä, Seppo & Kolehmainen, Leena. 2012. Weather and Language. *Language and Linguistics Compass* 6.6: 383-402.

Fattier, Dominique. 2013. Haitian Creole structure dataset. In Michaelis et al. (2013c).

Finney, Malcolm Awadajin. 2013. Krio structure dataset. In Michaelis et al. (2013c).

François, Alexandre. 2001. Contraintes de structures et liberté dans l'organisation du discours. Une description du mwotlap, langue océanienne du Vanuatu. PhD thesis Université Paris-IV Sorbonne.

Güldemann, Tom. 2010. Sprachraum and geography: Linguistic macro-areas in Africa. In Alfred Lameli, Roland Kehrein & Stefan Rabanus (eds.), *Language and space: An international handbook of linguistic variation*, vol. Vol. 2: *Language mapping*, 561–585, Maps 2901–2914. (HSK 30.2). Berlin, New York: De Gruyter Mouton.

Hagemeijer, Tjerk. 2013. Santome structure dataset. In Michaelis et al. (2013c).

Haspelmath, Martin. 2001. Non-canonical marking of core arguments in European languages. In Aikhenvald, Alexandra Y., Dixon, Robert M.W. and Onishi, Masayuki (eds). *Non-canonical Marking of Subjects and Objects*. Amsterdam; Philadelphia: John Benjamins. 53-83.

Haspelmath, Martin. 2005. Ditransitive Constructions: The Verb 'Give'. In: Haspelmath, Martin et al. (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 426-9.

Haspelmath, Martin & Michaelis, Susanne Maria. 2013. Ditransitive Constructions with 'give'. In Michaelis et al. (2013a), 236–39.

Haspelmath, Martin & Michaelis, Susanne Maria. 2014. Hybridization, social conditions and word order. Paper presented at the workshop "Grammatical hybridization and social conditions", MPI-EVA Leipzig 16-18 October 2014.

Huber, Magnus. 1999. *Ghanaian Pidgin English in its West African Context: A Sociohistorical and Structural Analysis*. Amsterdam: Benjamins.

Huber, Magnus. 2013. Ghanaian Pidgin English structure dataset. In Michaelis et al. (2013c).

Holm, John A. 1988. *Pidgins and creoles. Volume 1: Theory and structure*. Cambridge: Cambridge University Press.

Holm, John A. & Peter L. Patrick. 2007. *Comparative creole syntax: Parallel outlines of 18 creole grammars*. London: Battlebridge.

Intumbo, Incanha & Holm, John A. & Inverno, Liliana. 2013. Guinea-Bissau Kriyol structure dataset. In Michaelis et al. (2013c)

Jackendoff, Ray. 1983. *Semantics and Cognition*, Cambridge, Massachusetts: MIT press.

Keesing, Roger M. 1985. *Kwaio Grammar*. Canberra: The Australian National University.

Kofod, Frances. 1978. The Miriwung Language. A Phonological and Morphological Study. MA thesis. University of New England.

Kortmann, Bernd & Kerstin Lunkenheimer (eds.). 2013. *Electronic World Atlas of Varieties of English (eWAVE)*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://ewave-atlas.org/.

Koopman, Hilda. 1986. The Genesis of Haitian. In Muysken, Pieter & Smith, Norval (eds.), 231-258

Kouwenberg, Silvia. 2013. Papiamentu structure dataset. In Michaelis et al. (2013c).

Kroeger, Paul. 1993. *Phrase Structure and grammatical relations in Tagalog*. Leland Stanford Junior University.

Lefebvre, Claire. 1998. *Creole genesis and the acquisition of grammar*. Cambridge: Cambridge University Press.

Lefebvre, Claire. 2015. *Functional Categories in Three Atlantic Creoles: Saramaccan, Haitian and Papiamentu*, Amsterdam: Benjamins.

Lefebvre, Claire (ed.). 2011. *Creoles, their substrates, and language typology*. Amsterdam: Benjamins.

Lefebvre, Claire & Brousseau, Anne-Marie. 2002. *A Grammar of Fongbe*, Berlin: Mouton de Gruyter.

Lehmann, Christian. 1992. Yukatekische lokale Relatoren in typologischer Sicht. In: *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung (ZPSK)*, 45: 626-641.

Lim, Lisa & Umberto Ansaldo. 2013. Singlish structure dataset. In Michaelis et al. (2013c).

Madan, Arthur Cornwallis. 1902. *English-Swahili Dictionary*. 2nd edition. Oxford: Clarendon Press.s

Maurer, Philippe. 2013a. Angolar structure dataset. In Michaelis et al. (2013c).

Maurer, Philippe. 2013b. Batavia Creole structure dataset. In Michaelis et al. (2013c).

Maurer, Philippe. 2013c. Principense structure dataset. In Michaelis et al. (2013c).

McWhorter, John H. 1997. *Towards a new model of creole genesis*. New York: Peter Lang.

Michaelis, Susanne Maria. 2016. Avoiding bias in comparative creole studies: Stratification by lexifier and substrate. *Journal of language contact*.

Michaelis, Susanne. 2008. Valency patterns in Seychelles Creole: Where do they come from? In Michaelis, Susanne (ed.). *Roots of creole structures*: *Weighing the contribution of substrates and superstrates*. Amsterdam: Benjamins, 225-51.

Michaelis, Susanne Maria & the *APiCS* Consortium. 2013a. Raining constructions. In Michaelis et al. (eds.) 2013a, 256-59.

Michaelis, Susanne Maria & the *APiCS* Consortium. 2013b. Experiencer constructions with 'headache'. In Michaelis et al. (eds.). 2013a, 260-63.

Michaelis, Susanne Maria & the *APiCS* Consortium 2013c. Experiencer constructions with 'like'. In Michaelis et al. (eds.). 2013a, 264-67.

Michaelis, Susanne Maria & the *APiCS* Consortium 2013d. Experiencer constructions with 'fear. In Michaelis et al. (eds.). 2013a, 268-71.

Michaelis, Susanne Maria & the *APiCS* Consortium 2013e. Motion-to and motion-from. In Michaelis et al. (eds.). 2013a, 322-5.

Michaelis, Susanne Maria & Rosalie, Marcel. 2013. Seychelles Creole structure dataset. In Michaelis et al. (2013c).

Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.) 2013a. *The Atlas of Pidgin and Creole Language Structures*, Oxford: OUP.

Michaelis, Susanne Maria & Maurer, Philippe & Haspelmath, Martin & Huber, Magnus (eds.) 2013b. *The Survey of Pidgin and Creole Language*, 3 volumes, Oxford: OUP.

Michaelis, Susanne Maria, Philippe Maurer, Martin Haspelmath, and Magnus Huber (eds.) 2013c. *APiCS Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://apics-online.info/.

Moyse-Faurie, Claire. 2013. Xârâcùù Valency Patterns. In: Hartmann, Iren & Haspelmath, Martin & Taylor, Bradley (eds.) 2013. *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Mufwene, Salikoko S. 2001. *The ecology of language evolution*. Cambridge: Cambridge University Press.

Muysken, Pieter & Smith, Norval (eds.). 1986. *Substrata versus Universals in Creole Genesis*, Amsterdam: Benjamins.

Muysken, Pieter & Smith, Norval (eds.). 2015. *Surviving the middle passage: The West Africa-Surinam Sprachbund*. Berlin: DeGruyter.

Neumann-Holzschuh, Ingrid & Klingler, Thomas A. 2013. Louisiana Creole structure dataset. In Michaelis et al. (2013c).

Nizar, Milla. 2010. Dative subject constructions in South-Dravidian Languages. PhD thesis Berkeley.

Osumi, Midori. 1995. *Tinrin grammar*. Honolulu: University of Hawai'i Press.

Paauw, Scott. 2013. Ambon Malay structure dataset. In Michaelis et al. (2013c).

Parkvall, Mikael. 2000. *Out of Africa: African Influences in Atlantic Creoles*. London: Battlebridge.

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

Rivierre, Jean-Claude. 1980. *La Langue de Touho: Phonologie et grammaire du Cèmūhî (Nouvelle Calédonie)*. Paris: Peeters.

Schröder, Anne. 2013. Cameroon Pidgin English. In Michaelis et al. (2013c).

Schultze-Berndt, Eva 2013. Jaminjung Valency Patterns. In: Hartmann, Iren & Haspelmath, Martin & Taylor, Bradley (eds.) 2013. Valency Patterns Leipzig. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://valpal.info/languages/jaminjung/examples/33, Accessed on 2015-09-21)

Schultze-Berndt, Eva & Angelo, Denise. 2013. Kriol structure dataset. In Michaelis et al. (2013c).

Siegel, Jeff. 1999. Transfert constraints and substrate influence in Melanesian Pidgin. Journal of Pidgin and Creole Languages. 14: 1-44.

Siegel, Jeff. 2008. *The Emergence of pidgin and creole languages*. Oxford: Oxford University Press.

Sippola, Eeva. 2013. Ternate Chabacano structure dataset. In Michaelis et al. (2013c).

Slomanson, Peter. 2013. Sri Lankan Malay structure dataset. In Michaelis et al. (2013c).

Smith, Geoff P. & Siegel, Jeff. 2013a. Tok Pisin structure dataset. In Michaelis et al. (2013c).

Smith, Geoff P. & Siegel, Jeff. 2013b. Tok Pisin. In Michaelis et al. (2013b), 214-22.

Subbarao (eds.), *Non-nominative Subjects*, Volume 1. 227-244. Amsterdam: Benjamins.

Zhang, Guohua 2013. Mandarin Chinese Valency Patterns. In: Hartmann, Iren & Haspelmath, Martin & Taylor, Bradley (eds.) 2013. *Valency Patterns Leipzig*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

van den Berg, Margot & Bruyn, Adrienne. 2013. Early Sranan structure dataset. In Michaelis et al. (2013c).

Velupillai, Viveka. 2013. Hawai'i Creole structure dataset. In Michaelis et al. (2013c).

Vitale, Anthony. 1981. *Swahili Syntax*. Dordrecht: Foris.

Wälchli, Bernhard & Zúñiga, Fernando. 2006. Source-goal (in)difference and the

typology of motion events in the clause. *Sprachtypologie und Universalienforschung* (*STUF)* 59. 284-303.

Winford, Donald & Plag, Ingo. 2013. Sranan structure dataset. In Michaelis et al. (2013c).

Winford, Donald & Migge, Bettina. 2007. Substrate influence on the emergence of the TMA systems of the Surinamese creoles. *Journal of Pidgin and Creole Languages* 22. 73-99.