

Peter Bakker (Aarhus University)

Are non-European creole languages structurally distinct from European creoles?

Robust empirical research in recent years has shown that creoles are typologically distinct from non-creoles (Szmrecsany & Kortmann for English creoles, McWhorter 2005, Bakker et al 2011 for creoles in general). A number of objections have been raised against this claim, and one of them was the underrepresentation of non-European-based creoles in Bakker et al.'s worldwide samples. It is indeed the case that fewer than a handful non-European creoles were included in that study.

In this paper we will look at a number of languages of which it has been claimed that they are creoles based on non-European languages. We will evaluate the claims, to the extent that information was retrievable, about a selection of them, including: Rao, a possibly creolized Papuan language of the Sepik region; Yilan Creole Japanese of Taiwan; Grand Ronde Chinook Wawa, a nativized form of Chinook Jargon. We will compare the structural properties of these creoles with those often associated with creoles, and with a sample of the languages of the world. We will focus on a number of unusual features of these languages, such as pronominal case and SOV constituent order in Rao, negative affixation and verb-final order in Yilan creole. Such non-European creoles shed light on typological properties of creoles: they deviate from Atlantic and Pacific types of creoles, but share a significant part of properties with them as well.

Bakker, Peter, Aymeric Daval-Markussen, Mikael Parkvall & Ingo Plag. 2011. Creoles are typologically distinct from noncreoles. *Journal of Pidgin and Creole Languages* 26(1): 5-42.
McWhorter, John H. 2005. *Defining Creole*. New York/Oxford:Oxford University Press.
Szmrecsanyi, Benedikt & Bernd Kortmann. 2009. The morphosyntax of varieties of English worldwide: a quantitative perspective. *Lingua* 119(11): 1643-1663.

Subject null arguments in creole languages

Previously it was believed that creole languages obligatorily require overt subject pronouns - notwithstanding some exceptions such as null expletive subjects - as a result of the lack of verbal inflection in these languages. Even if this holds true for the great majority of creole languages (cf. APiCS), more recent studies have shown that this is not the case of many Asian and Indian Ocean creoles, in which pronominal subjects are frequently "omitted".

This study provides an account of subject pronoun usage in creoles focusing on a number of creoles that represent different areal patterns. Our sample includes the Atlantic creoles of San Andrés, Nicaragua, Berbice Dutch and Cape Verdean, the Indian Ocean creoles of Réunion and Mauritius, Diu Indo-Portuguese, and Philippine Creole Spanish.

Given the shallow chronological depth of these contact languages - which often involve lexifier languages with variable subject marking -, the creoles under survey offer an ideal corpus for studying the processes involved in the evolution of null subjects.

As Wratil (2011) has shown, the occurrence, distribution and development of null subject pronouns is not necessarily connected with characteristics of verbal inflection or with word order. By consequence, we turn to discourse as the main explanatory factor. We shall explore the discourse properties and other factors in verbal semantics that favor the occurrence of null subjects in our sample. We also use Travis and Torres Cacoullós' (2012; 2013a; 2013b) cross-linguistic generalizations based on constraints on variable subject realization, especially in cases not optimally catered for by the definitions presented in the APiCS database. Nevertheless, our results confirm that substrate/adstrate influence is the main factor which affects the expression of subjects in creoles.

Dryer, Matthew S.. 2011. Expression of Pronominal Subjects. In: Dryer, Matthew S. & Haspelmath, Martin (eds.) *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 101. Available online at <http://wals.info/chapter/101> Accessed on 2013-04-15.

Haspelmath, Martin and the APiCS Consortium. 2013. Expression of pronominal subjects. In: *The Atlas of Pidgin and Creole Language Structures*, ed. by Susanne Maria Michaelis, Philippe Maurer, Martin Haspelmath, and Magnus Huber. Oxford: Oxford University Press.

Lipski, John. 1995. The evolution of null subjects in Philippine Creole Spanish. 1994 Mid-America Linguistics Conference Papers, Volume II. Lawrence, Kansas: University of Kansas Linguistics Department. 387-401

Lipski, John. Null subjects in Romance-derived creoles: routes of evolution, Society for Pidgin and Creole Linguistics annual meeting, Los Angeles, January 8, 1999.

Torres Cacoullós, Rena and Catherine E. Travis. 2013a. Prosody, priming and particular constructions: The patterning of English first-person singular subject expression in conversation. *Journal of Pragmatics*. (available at: <http://dx.doi.org/10.1016/j.pragma.2013.08.003>)

Torres Cacoullós, Rena and Catherine E. Travis. 2013b. Exploring proposed cross-linguistic tendencies: A variationist approach to subject realization. Paper presented at the SKY symposium Subject: cognitive, typological and functional approaches Helsinki, September

12-13, 2013.

Wratil, Melanie. 2009. Null arguments and homonymy flight in the development of Creole pronoun systems. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS. www.hrelp.org/eprints/ldlt2_31.pdf

Wratil, Melanie. 2011. Uncovered pro – On the development and identification of null subjects. In Melanie Wratil & Peter Gallmann (eds.) *Null Pronouns. (Studies in Generative Grammar 106)*. Berlin: Walter de Gruyter. 99-140.

Wratil , Melani. 2009. Null arguments and homonymy flight in the development of Creole pronoun systems. In Peter K. Austin, Oliver Bond, Monik Charette, David Nathan & Peter Sells (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory 2*. London: SOAS. www.hrelp.org/eprints/ldlt2_31.pdf

Adrienne Bruyn (University of Utrecht)

Suriname versus the rest of the world: a comparison of functional elements based on the APiCS

Taking data from the APiCS database as a starting point, I will compare functional elements of the creole languages of Suriname with those occurring in other creole languages. I will take into account creoles with the same lexifier languages as the Suriname creoles, that is, English and also Portuguese and Dutch, both within the Atlantic area and elsewhere, and those with a similar substrate, in particular Haitian.

The functional elements to be discussed include aspect markers, complementizers and adpositions. With respect to the latter, the Suriname creoles exhibit a pattern that appears rather unique: based on the Gbe substrate, items derived from English and Portuguese are used in complex adpositional phrases. Not only are such complex adpositional phrases not attested in other English or Portuguese creoles, they are also absent from Haitian which has French as its lexifier but shares its substrate with the Suriname creoles.

Although the differences between the creoles of Suriname and those that developed elsewhere are not always as absolute as in the case of adpositions, there is diversity both in the selection of lexifier items to perform a particular function and in the nature of similar functional items across creoles. A comparison as outlined above may shed light on the nature and the respective roles of lexifier input, substrate influence and innovation in creole development.

By way of a methodological note, I will also pay attention to the issue of presentday data and the historical dimension.

Aymeric Daval-Markussen (Aarhus University)

Creole typology in the age of the APiCS

In recent years, the field of historical linguistics has witnessed a surge of research taking advantage of tools developed for the analysis of evolutionary biological phenomena (e.g. McMahon & McMahon 2003, Gray & Atkinson 2003). These tools have also been applied to language groups with a long history of contact (Dunn et al. 2005, 2008). More recently, contact languages have been the main focus of such studies (Bakker et al. 2011, Daval-Markussen 2011).

With the publication of the *Atlas of Pidgin and Creole Language Structures* (APiCS, Michaelis et al. 2013), an unprecedented database on contact varieties is now available. Conveniently, a number of features described in the APiCS overlap with those documented in the *World Atlas of Language Structures* (WALS, Haspelmath et al. 2008).

In this paper, we will use the joint datasets of the APiCS and the WALS in order to provide further empirical evidence supporting the claim that creoles constitute a synchronically distinguishable subgroup among the world's languages and therefore confirm the results provided in Bakker et al. (2011). Various phylogenetic and statistical methods will be employed to reach this goal with a considerable degree of certainty.

Bakker, Peter, Aymeric Daval-Markussen, Mikael Parkvall & Ingo Plag. 2011. Creoles are typologically distinct from non-creoles. In Parth Bhatt & Tonjes Veenstra (eds.), *Journal of Pidgin and Creole Languages*. 26(1): 5-42.

Dunn, Michael, Stephen C. Levinson, Eva Lindström, Ger Reesink & Angela Terrill. 2008. Structural Phylogeny in Historical Linguistics: Methodological Explorations Applied in Island Melanesia. *Language*. 84: 710-759.

Dunn, Michael, Angela Terrill, Ger Reesink, Robert A. Foley & Stephen C. Levinson. 2005. Structural phylogenetics and the reconstruction of ancient language history. *Science*. 309: 2072-2075.

Gray, Russell D. & Quentin D. Atkinson. 2003. Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature*. 426: 435-439.

Haspelmath, Martin, Matthew S. Dryer, David Gil, & Bernard Comrie (eds.). 2008. *The World Atlas of Linguistic Structures*. Munich: Max Planck Digital Library.

McMahon, April & Robert McMahon. 2003. Finding families: Quantitative methods in language classification. *Transactions of the Philological Society*. 101: 7-55.

Michaelis, Susanne, Philippe Maurer, Martin Haspelmath & Magnus Huber. 2013. *The Atlas of Pidgin and Creole Structures*. Oxford: Oxford University Press.

Michael Dunn (Max Planck Institute for Psycholinguistics, Nijmegen)

A time and place for languages: new perspectives from Bayesian Phylogeography

When were ancestral languages spoken, and where? Beyond the few cases where written texts of ancient languages have been preserved, it has proven extremely difficult to link reconstructed languages to concrete places and exact times. After an optimistic beginning in the 1950s, quantitative approaches to historical linguistics (glottochronology) suffered a collapse in confidence from which the discipline has only recently begun to emerge. New, phylogenetic approaches allow linguistics to answer old questions through new methods for inferring language family histories. And linguistic phylogenetic methods have also opened up new areas of enquiry, notably by using phylogenetic trees as a statistical 'scaffold' to test hypotheses about other historical process, ranging from population dispersal order (Gray and Jordan 2001) to evolutionary dependencies in structural change (Dunn et al. 2011).

Methodologically, linguistic phylogenetic approaches have been making great advances since the adoption of Bayesian phylogenetic inference methods. Bayesian phylogenetic inference combines more realistic models of language evolution with statistically sound methods of parameter estimation and model testing. The modeling methods include inferences about rate of change, and so an important early refinement to the methods was to allow external calibration points (a priori known dates for a subset of nodes on the tree) to be introduced, which in turn allow rates to be converted to dates.

Bayesian phylogenetic inference allows many different kinds of data to be incorporated, and one further refinement is to include spatial information about languages in the analysis. Models of spatial variation allow the known location of specified nodes of the tree to contribute information to the phylogenetic inference procedure, which in turn allows the spatial location of hitherto unspecified nodes to be inferred. These new methods of phylogenetic inference and the hypothesis testing techniques will be illustrated with case studies from a range of language families, including Indo-European, Aslian and Uralic.

Experiments with the basic vocabulary of the Gulf of Guinea Creoles

This paper sets out by measuring the orthographic and phonetic distance between the basic vocabulary (extended Swadesh list) in the four Portuguese-related Gulf of Guinea creoles: Angolar (AN), Santome (ST), Principense (PR) and Fa d'Ambô (FA). We expect that the differences between orthography and phonetics not to be significant, because the writing system is already based on phonetics and phonology. Using computational methods, we will attempt to reconstruct the basic vocabulary of the proto-language from which the four creoles descend.

Keywords: Gulf of Guinea creoles, lexical distance, lexical reconstruction

1. Comparing their orthography

We use dynamic programming to compute the edit-distance between two words. This is the minimum number of operations (deletion, substitution, insertion) required to align two words together and this is called the cost or the distance between the words. Note that we use a uniformized writing system for the four creoles. Average distance per word-pair We measured the distance between a Swadesh list of 216 terms and a reduced list of 163 terms without African items, which are especially common in Angolar. We found that the distance between languages is less when we used the reduced list and that AN, FA and PR creoles are closer to ST than to each other.

2. Comparing their phonetics

This time we used edit-distance and the same vocabularies (Swadesh and reduced) to compute phonetic distance but with a phonetic model and a larger set of operations (e.g. fortition, lenition, aphaeresis, prosthesis, apocope, etc.). Note that the differences between orthography and phonetics are not significant, because the writing system is already based on phonetics-phonology.

Average cost per word-pair Results are similar to orthographic distance, albeit distances are slightly more marked. We also found that AN-FA appears most distant orthographically and phonetically, while PR-ST appears least distant orthographically and phonetically.

3. Reconstructing proto-lexical items from the Swadesh list

We adapted a computer program (Oakes, 2000) which implements the "Comparative Method" described in Crowley and Bovern (1998).

Results We automatically reconstructed 79% of the Swadesh lexicon for the four creoles.

4. Future Work

- _ Examine the list of frequent sound changes to see if new regular patterns apply to Gulf of Guinea creoles.
- _ Use the distance measure to extract potential cognates from Gulf of Guinea creole corpora.
- _ Evaluate the list of proto-words we reconstructed.

Crowley, Terry and Claire Bovern. 1998. *An introduction to Historical Linguistics*. Oxford University Press.

Oakes, Michael P.. 2000. Computer estimation of vocabulary in a protolanguage from word lists in four daughter languages. *Journal of Quantitative Linguistics*, 7(3): 233–243.

Stéphane Goyette (Brandon University, Manitoba)

On the genesis of French and non-French Creoles of the Americas

From a typological point of view French Creoles of the Americas (hereinafter FCA) differ from English Creoles of the Americas and Papiamentu (hereinafter NFCA) in a number of ways. In sharp contrast to many if not most of its NFCA neighbors, no FCA makes phonemic use of tone, no FCA has any African-origin consonant phoneme, no FCA systematically eliminates word-final consonants, no FCA makes use of any borrowed African personal pronouns. More broadly, we may say that in all major respects, FCA's differ from NFCA's in showing considerably less African influence.

The goal of the proposed presentation is to seek to account for this set of differences between FCA's and NFCA's by proposing that whereas the latter derive from proto-pidgins spoken in West Africa (for English-based creoles, see McWhorter (1995), and for Papiamentu, Jacobs (2012)), the former must derive from a proto-pidgin originally spoken in the Americas. Parkvall (2000: 156) had already expressed the suspicion that a French proto-pidgin which arose in the Americas (unlike the ancestral proto-pidgins of NFCA's, with an alleged African place of birth) would explain the comparatively weak African influence upon FCA's. A likely candidate is the island of Saint-Kitts, which both the French and English settled in the seventeenth century, and which served as a springboard for the subsequent French and English colonization of other West Indian islands. Linguistic evidence in support of this hypothesis is the hitherto-unnoticed fact that a good many linguistic features of FCA's point to their common ancestor having been in contact with different varieties of English.

Jacobs, Bart. 2012. *Origins of a Creole: The History of Papiamentu and Its African Ties*. Berlin: Walter DeGruyter.

McWhorter, John. 1995. "Sisters under the skin: A case for genetic relationship between the Atlantic English-based creoles". *Journal of Pidgin and Creole Languages*: 10: 289-333.

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

Rebecca Grollemund, Simon Branford & Mark Pagel (University of Reading)

Bantu expansion follows the Savannah Corridor through the Equatorial rainforest

We present a phylogenetic classification of the Bantu languages based on new lexical data for 450 languages, using a novel Bayesian phylogenetic methods that allow for varying rates of lexical evolution. By applying a relaxed clock dating method we trace the Bantu expansion, finding the major migration routes and date estimates for all nodes. Our results imply a different major migration route than has been proposed in several recent studies.

The initial migration, starting from the Bantu homeland in the Mbam region (Cameroon), agrees with previous work. The model infers a main migration wave, beginning around 4500 BP, south towards Cameroon and Gabon and the rainforest. At around 2500 BP we find evidence for a big split of the Bantu people. This key aspect of the Bantu expansion occurs when the ancestors of the present day Western Bantu migrated south from Northern Congo through the rainforest following a newly opened savannah corridor. The other group, the ancestors of the present day Eastern Bantu, continued along the north and east edges of the rainforest and settled the Eastern and Southern Bantu areas.

Our results differ from recently published work by showing that Bantu expansion was influenced by the environmental and climatic conditions. Indeed, our results fit nicely with the current hypotheses about the climate-induced vegetation change that occurred 2500 years ago and which opened savannah “corridors” through the equatorial forest (Maley, 2001; Bostoen and al., 2013). The climatic conditions have facilitated the Bantu expansion and influenced the Bantu population movements.

Bostoen, K., Grollemund, R. & Koni Muluwa, J. 2013. Climate-induced vegetation dynamics and the Bantu Expansion: Evidence from Bantu names for pioneer trees (*Elaeis guineensis*, *Canarium schweinfurthii*, and *Musanga cecropioides*). *C. R. Geoscience*.

Maley, J. 2001. La destruction catastrophique des forêts d’Afrique centrale survenue il y a environ 2500 ans exerce encore une influence majeure sur la répartition actuelle des formations végétales. *Systematics and Geography of Plants* 71: 777-796.

Basic Word Order and Language Contact

It is commonly suspected that SVO word order is associated with high-contact languages, after the observation that a number of high-contact languages display this word order. However, if high-contact was indeed the driving factor, we would expect to see SVO word order to systematically appear in pockets throughout the world's languages, since languages are in contact everywhere, and high-contact scenarios occur, and can be assumed to have occurred in past, in pockets everywhere on the planet. However, informal inspection of the occurrence of SVO word order instead reveals a high concentration of SVO along certain lineages and certain areas, not *prima facie* associated higher contact than areal concentration of other word orders. This calls for a systematic investigation.

To this end we will perform three tests regarding basic word order, as follows. Using a recently compiled database of basic word order for over 4600 languages (1) we perform the following tests:

_ **Word Order Borrowing:** Using known family relationships and geographical coordinates, we can do perform maximum-parsimony analysis of the attested word order data infer 1) likely word orders of proto-languages at every stage, and 2) likely word order borrowings. This allows us to find out whether SVO word order is borrowed more often than other word orders.

_ **Word Order Borrowing:** For most of the 4600 languages, the practically usable sociohistoric data are speaker number estimates. Using these and the analysis in the previous point we can investigate whether a smaller language is more likely to have borrowed from a larger neighbour (i.e., by direct influence) or vice versa (i.e., by substrate effects).

_ **Contact-Favoured Word Order:** Suppose A is the parent of a language B which neighbours another language C. If there is a word order favoured by contact per se (i.e., without any of the involved languages actually displaying it, which is required to infer direct borrowing), it should be overrepresented regardless of the configuration of A and C.

¹This database is exhaustive in the sense that there is no published data on the basic word order of the remaining languages.

Vittorio Loreto (Sapienza University of Rome)

Modelling the emergence of creoles languages

Creole languages are nonstandard hybrid idioms which originated in the contacts of European colonists with typically the slaves they brought to their plantation settlement colonies. Some of the best known include Gullah and Jamaican Creole (lexified by, i.e., deriving most of its vocabulary from, English) as well as Haitian and Louisiana Creoles (lexified by French). Although most of the scholarship on these new vernaculars has been about their emergence, no attempt has been made to model the dynamical aspect of this evolutionary process, which can help resolve some of the controversies about what is involved in Creole-formation. In this talk I'll present a dynamical process that mimics the process of Creole formation in American contact ecologies involving European colonists and African slaves in the Eighteenth century. Inspired to the Naming Game, our modeling scheme incorporates demographic information about the composition of the colonial position during that period, which included Europeans, Mulattos/Creoles (both locally born), and Bozal (African-born) slaves. I'll show how this sole information allows to discriminate regions/territories that produced modern creole idioms from those that did not with a surprising accuracy. From this perspective the presented modeling scheme may turn out to be a useful tool to shed light on the formation of creole languages as well as to test specific hypotheses as they apply to the relevant ecologies. The basic algorithm is, in fact, flexible enough to accommodate additional relevant factors to explain specific spatial heterogeneities in creole formation. Finally the generality of the approach could provide valuable insights for further studies related to the emergence of languages in contact linguistic ecologies along as for the longstanding problem of the emergence and evolution of Language.

Susanne Michaelis (Max Planck Institute for Evolutionary Anthropology, Leipzig)

Sampling in contact linguistics: What is a typical creole feature?

Creolists sometimes claim that there are typological generalizations across contact languages (e.g. Bickerton 1981, McWhorter 2001, Bakker et al. 2011). But they have generally overlooked that cross-creole generalizations require representative samples, especially when working quantitatively. Sampling for genealogical and areal control has been a much discussed topic within world-wide typology, but not yet in comparative creolistics.

In all available comparative creoles studies, European-based Atlantic creoles are strongly overrepresented, so that typical features of these languages are taken as typical creole features, e.g. serial verbs, or 3PL pronouns as nominal plural markers. But many of these creoles have the same genealogical/areal profile, i.e. European (lexifier) + Macro-Sudan (substrate(s)). I therefore propose a sampling method that controls for genealogical/areal relatedness of both the substrate(s) and the lexifier, which I call "bi-clan" control. A "clan" is a language family or linguistic area, and a "bi-clan" is a combination of a lexifier clan and a substrate clan. For example, Atlantic English (pidgin)creoles belong to the bi-clan English + Macro-Sudan, Oceanic English (pidgin)creoles belong to the bi-clan English + Oceanic.

Using APiCS data, I will show that there are implications of bi-clan sampling for the notion of so-called "typical creole features". For instance, directional serial verb constructions (with 'come' and 'go') are present in 52% of the APiCS languages, but only in 35% of the bi-clans, simply because of the oversampling of the bi-clans English/French/Portuguese + Macro-Sudan. On the other hand, features which have been classified as rare creole features (e.g. inclusive/exclusive distinction in pronouns, purely aspectual tense-aspect systems) may simply seem to be rare because of the underrepresentation of the non-Atlantic creoles. For instance, 13% of the APiCS languages have a purely aspectual tense-aspect system, whereas 75% show a mixed aspectual-temporal system. When represented in terms of bi-clans, the figures change: 20% have a purely aspectual system whereas only 46% show a mixed system. Therefore, bi-clans give a much more realistic picture of the structural properties of pidgins/creoles, and any notion of "typical creole feature" must be checked against a genealogically/areally representative sample of contact languages.

Erich Round (University of Queensland)

Dataset design processes need to be scientifically reported: The sensitivity of Bayesian clustering and ‘researcher degrees of freedom’

Modern computational methods open up new directions of enquiry utilizing cross-linguistic datasets whose size would overwhelm traditional pencil-and-paper analyses. Naturally however, questions arise when using techniques not originally designed for linguistic research, as to how performance might be affected by properties of linguistic datasets, whose principles of design have yet to alter much from the paper-and-pencil era. Issues such as non-independence of variables may be innocuous in some cases (Pagel & Meade 2006), but precisely when and whether this generalizes requires ongoing investigation.

In a series of experiments we examine the performance of the Bayesian clustering algorithm STRUCTURE (Pritchard et al. 2000) by making controlled variations to <10% of a large cross-linguistic dataset of typological variables (Reesink et al. 2009—121 languages; 155 binary variables; 88% of cells complete). We find that in terms of both the number of clusters inferred, and the posterior probability of those inferences, STRUCTURE is remarkably sensitive to the inclusion/exclusion of variables which have:

- i. A high proportion of the same value across languages;
- ii. A high number of missing values;
- iii. A tendency to pair-wise implication with another variable (e.g. A=1 very often entails B=0);
- iv. A high pair-wise correlation with another variable

This may carry consequences for the methodology of modern linguistic dataset construction, since a typologist’s decision over whether or not to build variables A, B or C into a dataset may hinge precisely on issues such as (i–iv), which can then influence the final statistical result — a manifestation of the “researcher degrees of freedom” problem (Simmons et al. 2011). Thus, to promote transparency and aid evaluation of research, we advocate that when datasets are designed, researchers explicitly document and publish the accompanying decision-making process, since as our results indicate, such decisions could bear significantly upon eventual results, before language coding even begins.

Pagel M., Meade A. (2006) Estimating rates of lexical replacement on phylogenetic trees of languages. In: Forster P, Renfrew C, editors, *Phylogenetic methods and the prehistory of languages*, McDonald Institute for Archaeological Research: UK, McDonald Institute Monographs. 173–182.

Pritchard J., Stephens M., Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Reesink G., Singer R., Dunn M. (2009) Explaining the linguistic diversity of Sahul using population models. *PLoS Biology* 7, e1000241.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366.

Eeva Sippola & Abigail Tiny (Aarhus University & Lisbon University)

Noun phrases in Iberian-lexifier creoles

This study provides an account of nominal syntax in Iberian-lexifier creoles from several areal clusters with different substrate languages. In creole studies, properties of noun phrases have received considerable attention (e.g. Baptista & Guerón 2007, Bickerton 1981, Bruyn 1995, Holm 1990), but to date, there have been no studies with a comparative approach on the Spanish- and Portuguese-lexifier creoles. Based on a cross-linguistic comparison with phylogenetic tools, the present study aims to reveal the shared features in the area of nominal syntax and the role of the noun phrase in the typological groupings of these creoles.

Our sample includes both Atlantic and Asian varieties and their lexifier languages Portuguese and Spanish. A phylogenetic network analysis of 22 features of nominal syntax from the APiCS database reveals that while Ibero-Asian creoles cluster as expected, Ibero-Atlantic creoles show considerable differences: Palenquero, Papiamentu, Guinea Bissau creole and the Gulf of Guinea creoles depart from the established areal clusters and proposed scenarios of evolution.

In the presentation, we will identify features responsible for the divergence and connect them with substrate influence and the evolutionary histories of these creoles based on complementary data from the project *the Origins and Development of Creole Societies in the Gulf of Guinea* (University of Lisbon). In addition, we will address the implications for phylogenetic network analysis based on typological data.

The findings of the study contribute to research of the Gulf of Guinea creoles in particular and typological approaches to Iberian-lexifier creoles in general.

William H. Sutherland (University of Cambridge)

Global distribution and changes of languages and biodiversity

William J. Sutherland¹, Tatsuya Amano¹, Brody Sandel², Heidi Eager^{3,4}, Edouard Bulteau⁵, Jens-Christian Svenning², Bo Dalsgaard⁶, Carsten Rahbek⁶, and Richard G. Davies⁷

¹ Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK.

² Ecoinformatics and Biodiversity Group, Department of Bioscience, Aarhus University, Aarhus DK-8000 C, Denmark.

³ Research Laboratory for Archaeology and the History of Art, School of Archaeology, University of Oxford, Oxford OX1 2HU, UK.

⁴ Department of Ecology and Evolutionary Biology, Cornell University, Corson Hall, Ithaca, NY 14853-2701, USA.

⁵ Ecole Polytechnique, Route de Saclay, 91120 Palaiseau, France.

⁶ Center for Macroecology, Evolution, and Climate, University of Copenhagen, Universitetsparken 15, DK-2100 Copenhagen, Denmark.

⁷ School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

There is a considerable body of research examining the pattern of diversity of species, examining the patterns of decline and the causes of decline. A comparable approach can be adopted for examining languages. The means of classifying the threats to species can also be applied to languages to provide a consistent measure. We will describe how these two subjects have much in common and show similar patterns of richness with similar correlates, but some key differences. We will describe two major analyses: one comparing different countries and one looking at global distributions across a grid.

Kees Versteegh (Radboud Universiteit Nijmegen)

The expression of future tense in Arabic pidgins and creoles

According to the model proposed by Bybee, Perkins and Pagliuca (1994), expression of the future develops along certain pathways of grammaticization, the most important of which start with either the expression of obligation or of desire, and end in the expression of intention > prediction. In this paper, I propose to analyze expressions of future tense in four different varieties of Arabic: South Egyptian Arabic/Sudanese Arabic as an example of a native variety; Pidgin Madame as an example of an early pidgin (Bizri 2010); Juba Arabic as an example of a stable pidgin (Tosco 1995); and (Ki-)Nubi as an example of an Arabic creole (Wellens 2005; Luffin 2005). Text collections of Pidgin Madame and Ki-Nubi have been published; for Sudanese Arabic and Juba Arabic, a corpus is available in text repositories on the internet. References to the future in the corpus will be analyzed semantically in order to verify whether the differences between the varieties correspond to the proposed pathways and which pathways can be distinguished in the development of Arabic.

Bizri, Fida. 2010. *Pidgin Madame: Une grammaire de la servitude*. Paris: Geuthner.

Bybee, Joan, Revere Perkins and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago and London: University of Chicago Press.

Luffin, Xavier. 2005. *Un créole arabe: Le kinubi de Mombasa, Kenya*. Munich: Lincom Europa.

Tosco, Mauro. 1995. "A pidgin verbal system: The case of Juba Arabic". *Anthropological Linguistics* 37.423-459.

Wellens, Ineke. 2005. *The Nubi language of Uganda: An Arabic creole in Africa*. Leiden: E.J. Brill.

Søren Wichmann (Max Planck Institute for Evolutionary Anthropology, Leipzig)

Worldwide patterns of language migration

In this paper techniques of dating proto-languages (Holman et al. 2011) and inferring their spatial location (Wichmann et al. 2010) are combined and applied to a dataset of over 4400 languages (Wichmann et al. 2013) in order to estimate patterns of language migration across the world and during different time periods. Different factors influencing migration rates are investigated, including the landscape, subsistence patterns, world regions, and time periods. Only some features of the landscape seem to have a decisive influence. Subsistence patterns are important, but apparently only in very recent times. Different world regions tend to exhibit similar patterns, although migration within Sahul proceeds at a slower pace than elsewhere. Finally, the overall trends in the increase and decrease of migration rates over time will be presented. The findings will be discussed in relation to recent studies of migration rates drawing upon archaeological and genetic data, which show some remarkable similarities.