Impact of climatic and ecological contexts on sociolinguistic factors among Bantu populations

Christophe Coupé[1,2] & Jean-Marie Hombert[1]

[1]Laboratoire Dynamique du Langage, CNRS - Université de Lyon, [2]Institut Rhône-Alpin des systèmes complexes

Background

- The expansion of Bantu farmers 3,000 years ago may be linked to climatic changes during the Holocene period (Schwartz, 1992)
- Questions remain about where Bantu populations came from and how they spread, especially when they faced the equatorial forest (Vansina, 1992)
- Geographic variations in terms of diversity of the Bantu languages; does highest linguitic density mean point of origin (Ehret, 2001)?
- Previous studies linking linguistic diversity to ecological factors, e.g. the length of the growing season (Nettle, 1996; Nettle, 1998; Jacquesson, 2001)

Goal

- To better understand the relationship between i) social and sociolinguistic factors and ii) African ecological contexts
- To later use this relationship to derive past human migrations from descriptions of their ecological contexts (e.g. with paleoclimatic models)

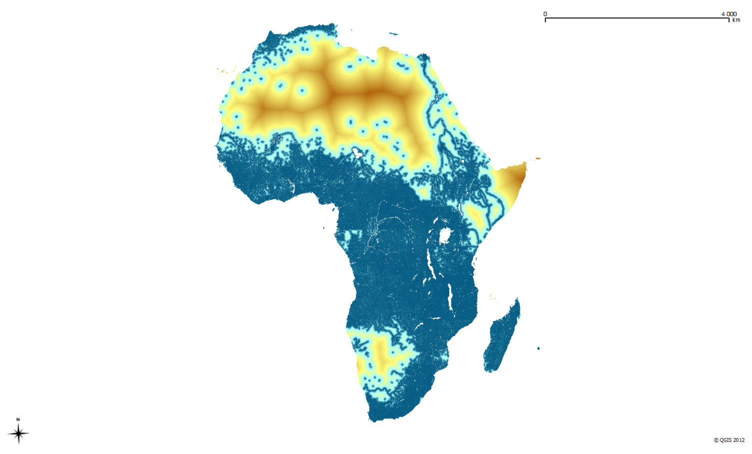
Methodology and tools

Methodology

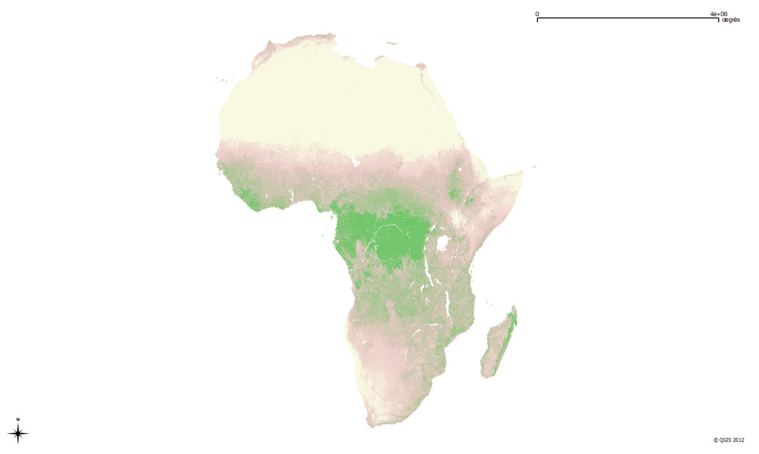
- To assemble a number of meaningful variables - elevation, vegetation etc. - describing African ecological contexts with high spatial resolution
- To conduct statistical analyses of the correlations between these variables and i) population density, ii) linguistic diversity

Datasets

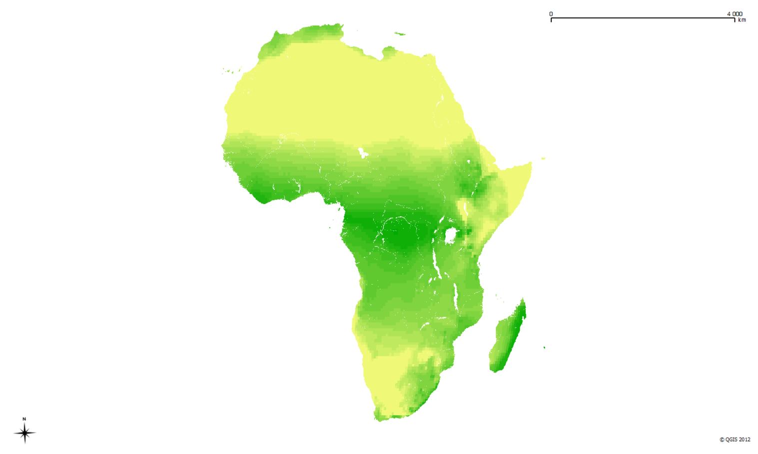
Elevation & rugosity



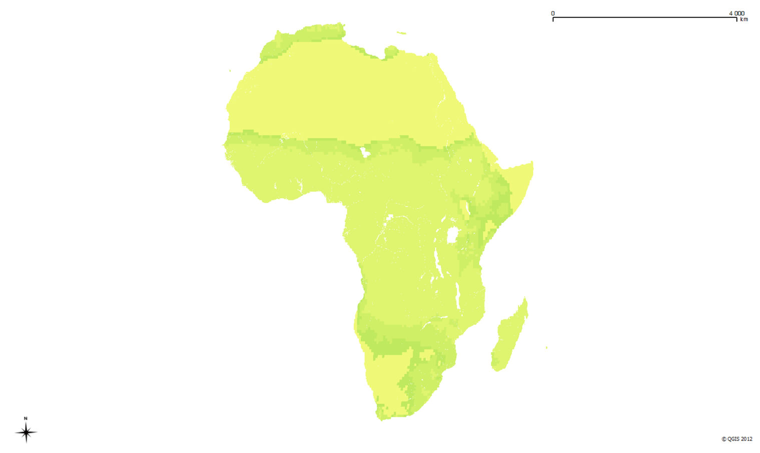
Distance to fresh water



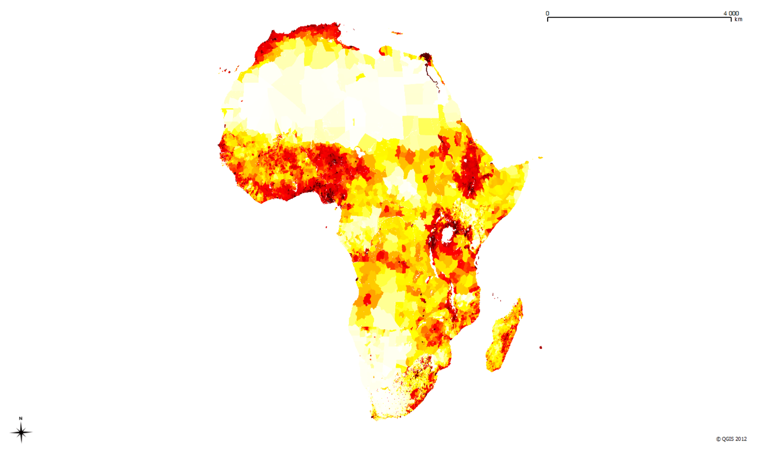
% of herbaceous plants and trees



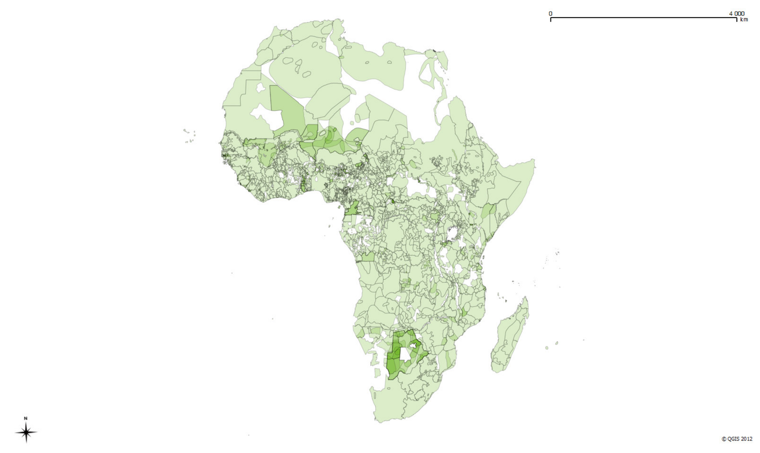
Length of the growing season (LGP)



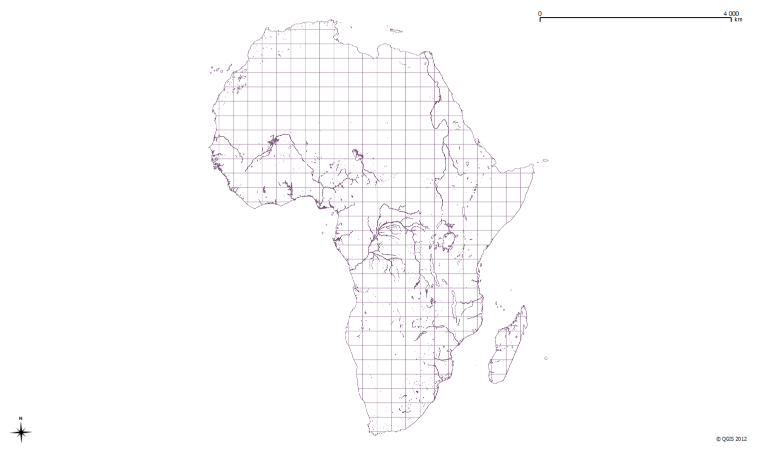
Inter-annual CV of the LGP



Population counts / density



Languages area & number of speakers



Units of statistical analysis

Statistical analyses

- Tools: *Quantum GIS*, *Excel*, R & *SPSS 19*
- Statistical units: 368 cells covering African lands (from 10,000 to 90,000 km²)
- *Boxplot* transformations for elv., rug., dist. to water, pop. dens. & ling. variables
- Significant colinearity between the ecological variables

→ difficult to identify 'primary' vs. 'secondary' causal factors

Results

Predicting the density of human populations

- Stepwise linear regression with POPULATION DENSITY as dependent variable

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate
1 (a)	.670	.449	.447	1.8012232
2 (b)	.718	.515	.512	1.6919270
3 (c)	.739	.547	.543	1.6378955
4 (d)	.785	.616	.612	1.5101667
5 (e)	.795	.632	.627	1.4802679
6 (f)	.801	.641	.635	1.4633308
7 (g)	.817	.667	.661	1.4112213

a. Predictors: (Constant), Distance to water
b. Predictors: (Constant), Distance to water, % herbaceous
c. Predictors: (Constant), Distance to water, % herbaceous, Elevation
d. Predictors: (Constant), Distance to water, % herbaceous, Elevation, Rugosity
e. Predictors: (Constant), Distance to water, % herbaceous, Elevation, Rugosity, CV LGP
f. Predictors: (Constant), Distance to water, % herbaceous, Elevation, Rugosity, CV LGP, % bare
g. Predictors: (Constant), Distance to water, % herbaceous, Elevation, Rugosity, CV LGP, % bare, LGP

Model 7: Coefficients					
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	sig
(Constant)	3.638	2.260		1.610	.108
Distance to water	-1.788	.345	-.383	-5.179	.000
% herbaceous	.080	.010	1.054	8.359	.000
Elevation	-.074	.009	-.301	-8.312	.000
Rugosity	.500	.085	.249	5.886	.000
CV LGP	.677	.155	.194	4.379	.000
% bare	.073	.012	1.281	6.188	.000
LGP	.342	.064	.593	5.306	.000

→ The density of population can be predicted with a high degree of accuracy (R > *0.8*) by ecological factors

Predicting the number of languages in a tile and language area

- Stepwise linear regression with NUMBER OF LANGUAGES as dependent variable

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate
1 (a)	.723	.523	.522	.6903445
2 (b)	.734	.538	.536	.6799944
3 (c)	.741	.550	.546	.6725704
4 (d)	.746	.556	.551	.6687364
5 (e)	.750	.562	.556	.6648291
6 (f)	.753	.567	.560	.6621412

a. Predictors: (Constant), % bare
b. Predictors: (Constant), % bare, Rugosity
c. Predictors: (Constant), % bare, Rugosity, Pop density
d. Predictors: (Constant), % bare, Rugosity, Pop density, CV LGP
e. Predictors: (Constant), % bare, Rugosity, Pop density, CV LGP, Elevation
f. Predictors: (Constant), % bare, Rugosity, Pop density, CV LGP, Elevation, LGP

Model 6: Coefficients					
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	sig
(Constant)	2.652	.299		8.863	.000
% bare	-.013	.002	-.561	-6.071	.000
Rugosity	-.216	.040	-.261	-5.438	.000
Pop density	.094	.022	.229	4.254	.000
CV LGP	-.145	.073	-.101	-1.983	.048
Elevation	.012	.005	.119	2.658	.008
LGP	.036	.018	.153	1.986	.048

- Stepwise linear regression with LANGUAGE AREA as dependent variable

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate
1 (a)	.786	.619	.618	1.9980498
2 (b)	.802	.643	.641	1.9356130
3 (c)	.809	.655	.652	1.9062422
4 (d)	.819	.670	.666	1.8659055
5 (e)	.821	.674	.670	1.8560117

a. Predictors: (Constant), % bare
b. Predictors: (Constant), % bare, LGP
c. Predictors: (Constant), % bare, LGP, Pop density
d. Predictors: (Constant), % bare, LGP, Pop density, CV LGP
e. Predictors: (Constant), % bare, LGP, Pop density, CV LGP, Elevation

Model 5: Coefficients					
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta	t	sig
(Constant)	13.893	.834		16.649	.000
% bare	.048	.006	.633	8.160	.000
LGP	-.154	.049	-.201	-3.139	.002
Pop density	-.218	.055	-.163	-3.943	.000
CV LGP	.794	.203	.171	3.909	.000
Elevation	.023	.010	.070	2.209	.028

→ The number of languages in a cell and their area can be predicted with a good degree of accuracy by ecological factors (R > 0.75 & R > 0.8 resp.)

- The number of languages at the putative origin of Bantu migrations is higher than what the regression predicts: density not fully explained by ecology?
- The number of speakers can be predicted with a fair degree of accuracy by ecological factors (R =0.582) (*not shown*)

Conclusions & Perspectives

- Tight relationships between ecological factors and social/sociolinguistic variables

→ Given 'layered' paleo-environmental data, possibility to reconstruct migratory paths (assuming a time-enduring relationship to ecological contexts)

- Define better measure of linguistic diversity, e.g. including the level of families and/or typological features
- Better assess the putative origin of Bantu migrations

References

1. Amante, C. & Eakins, B. W. (2009). ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis. *NOAA Technical Memorandum NESDIS NGDC-24.*
2. Center for International Earth Science Information Network (CIESIN), Columbia University; and Centro Internacional de Agricultura Tropical (CIAT). 2005. Gridded Population of the World Version 3 (GPWv3): Population Density Grids. Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University. Available at http://sedac.ciesin.columbia.edu/gpw. (2011).
3. Ehret, C. (2001). Bantu Expansions: Re-Envisioning a Central Problem of Early African History. *The Int. Journal of African Historical Studies* 34(1): 5-41.
4. Food and Agriculture Organization & International Institute for Applied Systems Analysis (2000). Length of growing period, 1901-1996. Global agro-ecological zones. In FAO & IIASA, 2007, *Mapping biophysical factors that influence agricultural production and rural vulnerability* (H. von Velthuizen et al.)
5. Food and Agriculture Organization & International Institute for Applied Systems Analysis (2000). Coefficient of variation of length of growing period, 1901-1996. Global agro-ecological zones. In FAO & IIASA, 2007, *Mapping biophysical factors that influence agricultural production and rural vulnerability* (H. von Velthuizen et al.)
6. GEODAS Grid Translator - Design-a-Grid. http://www.ngdc.noaa.gov/mgg/gdas/gd_designagrid.html
7. Global Mapping International & SIL International (2012). *World Language Mapping System. Language area and point data for Geographic Information Systems (GIS).* http://www.worldgeodatasets.com/language/
8. Hansen, M., R. DeFries, J.R. Townshend, M. Carroll, C. Dimiceli, and R. Sohlberg (2003), Vegetation Continuous Fields MOD44B, 2001 Percent Tree Cover, Collection 3, University of Maryland, College Park, Maryland, 2001.
9. Jacquesson, F. (2001). Pour une linguistique des quasi-déserts. In A.M. Loffler-Laurian (ed.), *Etude de linguistique générale et contrastive. Hommage à Jean Perrot.* Paris : Centre de Recherche sur les Langues et les Sociétés, pp. 199-216.
10. Kummu, M., de Moel, H., Ward, P. J., Varis, O. (2011) How Close Do We Live to Water? A Global Analysis of Population Distance to Freshwater Bodies. PLoS ONE 6(6): e20578. doi:10.1371/journal.pone.0020578
11. Lewis, M. P. (ed.), 2009. *Ethnologue: Languages of the World, Sixteenth edition.* Dallas, Tex.: SIL International. http://www.ethnologue.com/.
12. Nettle, D. (1996). Language diversity in West Africa: An ecological approach. *Journal of Anthropological Archaeology* 15:403-438.
13. Nettle, D. (1998). Explaining global patterns of language diversity. *Journal of Anthropological Archaeology* 17: 354-74..
14. Schwartz, D. (1992). Assèchement climatique vers 3 000 B.P. et expansion Bantu en Afrique centrale atlantique : quelques réflexions. *Bulletin de la Société géologique de France* 163(3) : 353-361.
15. Vansina, J. (1995). New Linguistic Evidence and 'the Bantu expansion'. *The Journal of African History* 36(2): 173-195.

Financial support: ANR CLHASS; ASLAN Laboratory of Excellence

Contact: Christophe COUPE, christophe.coupe@ish-lyon.cnrs.fr