

Uncovering lost paths in the Congo rainforest: A new, comprehensive phylogeny of West-Coastal and Central-Western Bantu

Sara Pacchiarotti¹, Natalia Chousou-Polydouri², Jean-Pierre Donzo^{1,3}, Guy Kouarata¹, Lorenzo Maselli^{1,4}, and Koen Bostoen¹

¹UGent Centre for Bantu Studies, Ghent University, Belgium

²Department of Comparative Language Science, University of Zurich, Switzerland

³ISP Gombe, Kinshasa, DRC

⁴Service de Métrologie et Sciences du Langage, Laboratoire de Phonétique, Université Mons, Belgium

The Bantu Expansion is the initial spread of the Bantu languages and the communities speaking them over large parts of Central, Eastern, and Southern Africa from a homeland located in the borderland between present-day Nigeria and Cameroon. This major linguistic, cultural, and demographic process in Late Holocene Africa stands out in three respects: its vastness, its rapidness, and its predominantly longitudinal orientation (Bostoen 2018). A central question about this expansion is whether and how the first Bantu-speaking populations migrated through and settled in the Congo rainforest. While previous studies suggest that the movement of Bantu-speaking people might have been favored by a climate-induced forest reduction in the Sangha River Interval (SRI) around 2500 BP (Bostoen *et al.* 2015; Grollemund *et al.* 2015), a recent study suggests that this migration through the rainforest happened well before that period and was not facilitated by SRI savanna corridors (Koile *et al.* 2022).

Although the specific quantitative methods underlying their conclusions differ, Grollemund *et al.* (2015) and Koile *et al.* (2022) are not only both phylogenetic studies, but they are also based on exactly the same datasets of basic vocabulary, the same cognacy judgments and the same underrepresented and unbalanced sample of rainforest Bantu languages. What is more, recent scholarship has seriously challenged the idea that phylogenies based on modern Bantu languages may directly reflect the initial migration of Bantu speech communities (Bostoen 2018; Gunnink *et al.* 2022; Bostoen *et al.* forthcoming). Seidensticker *et al.* (2021) argue that a population collapse hit the entire Congo rainforest ~1,500 BP, which probably led to the extinction of many ancestral lineages of Bantu languages before the area was recolonized by Bantu speakers from ~1,000 BP onwards. Hence, many of the Bantu languages currently spoken in the Congo rainforest may have an ancestry there that is more than a millennium younger than previously assumed.

In order to shed new light on the initial expansion of Bantu languages through the Congo rainforest and how its signal in lexicon-based phylogenies got possibly blurred by language death and spread-over-spread-events, we present in this talk the preliminary results of a new comprehensive lexicon-based phylogeny. This on-going study focuses on two clades of rainforest Bantu languages, i.e., Central-Western (CWB) and West-Western or West-Coastal (WCB) Bantu. These two groups display different topologies in the phylogenies of Grollemund *et al.* (2015) and Koile *et al.* (2022). While in the first, WCB branches off after CWB, in the second CWB and WCB are parallel branches. Moreover, portions of what is CWB in Grollemund *et al.* (2015) cluster more closely with WCB in Koile *et al.* (2022).

While Grollemund *et al.* (2015), the most comprehensive phylogeny of the Bantu languages to date, includes 424 doculects for the entire family, our new phylogeny includes more than 350 varieties for two branches, CWB and WCB. Featuring many varieties spoken in the DRC that were never documented before (Kouarata *et al.* forthcoming), it closely reflects modern-day Bantu language diversity within the Congo rainforest. The vast majority of our data come from first-hand fieldwork and second-hand specialized sources other than Bastin *et al.* (1999). Cognacy judgments are performed by relying on a profound knowledge of the historical phonology of the languages in question (Rottland 1977; Koni Muluwa & Bostoen 2012; Bostoen & Koni Muluwa 2014; Donzo 2015; Pacchiarotti & Bostoen 2020; Pacchiarotti & Bostoen 2021; Pacchiarotti & Bostoen 2022). We use Lexedata (Kaiping *et al.* 2022) as a toolbox to edit and annotate our lexical dataset and MrBayes 3.2 (Huelsenbeck & Ronquist 2001; Ronquist *et al.* 2012) to produce a bayesian phylogeny, which is an important first step to refine our understanding of the layered history of rainforest Bantu languages. Additionally, our preliminary results will serve as the basis to develop and subsequently test new hypotheses regarding the colonization and recolonization of the Congo rainforest by Bantu-speaking peoples.

References

- Bastin, Y., A. Coupez and M. Mann. 1999. *Continuity and Divergence in the Bantu Languages: Perspectives from a Lexicostatistic Study*. Tervuren: Royal Museum for Central Africa.
- Bostoen, K. 2018. The Bantu expansion. In T. Spear (ed.), *Oxford Research Encyclopedia of African History*. Oxford: Oxford University Press, Oxford Research Encyclopedias Online <https://oxfordre.com/africanhistory/view/10.1093/acrefore/9780190277734.001.0001/acrefore-9780190277734-e-191>.
- Bostoen, K., B. Clist, C. Doumenge, R. Grollemund, J.-M. Hombert, J. Koni Muluwa and J. Maley. 2015. Middle to late Holocene paleoclimatic change and the early Bantu expansion in the rain forests of West Central-Africa. *Current Anthropology* 56: 354-384.
- Bostoen, K. and J. Koni Muluwa. 2014. Umlaut in the Bantu B70/80 languages of the Kwilu (DRC). *Transactions of the Philological Society* 112: 209-230.
- Bostoen, K., S. Pacchiarotti and W. Hubau. forthcoming. A la recherche des sentiers perdus du passé en forêt: Reflexions sur l'effondrement de la population en Afrique centrale il y a 1500 ans. *Bulletin des Séances de l'Académie Royale des Sciences d'Outre-Mer*.
- Donzo, J.-P. 2015. *Langues bantoues de l'entre Congo-Ubangi (RD Congo): documentation, reconstruction, classification et contacts avec les langues oubanguiennes*. Gand: Université de Gand (UGent), thèse de doctorat.
- Grollemund, R., S. Branford, K. Bostoen, A. Meade, C. Venditti and M. Pagel. 2015. Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America* 112: 13296-13301.
- Gunnink, H., N. Chousou-Polydouri and K. Bostoen. 2022. Divergence and contact in Southern Bantu language and population history: A new phylogeny in cross-disciplinary perspective [07 Nov. 2022]. *Language Dynamics and Change*: 1-58.
- Huelsenbeck, J. P., & Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics*, 17(8), 754–755.
- Kaiping, G., M.S. Steiger and N. Chousou-Polydouri. 2022. Lexedata: A toolbox to edit CLDF lexical datasets. *Journal of Open Source Software* 7: 4140.
- Koile, E., S.J. Greenhill, D.E. Blasi, R. Bouckaert and R.D. Gray. 2022. Phylogeographic analysis of the Bantu language expansion supports a rainforest route. *Proceedings of the National Academy of Sciences* 119: e2112853119.
- Koni Muluwa, J. and K. Bostoen. 2012. La diphtongaison dans les langues bantu B70-80 (Bandundu, RDC): typologie et classification historique. *Africana Linguistica* 18: 355-386.
- Kouarata, G.N., S. Pacchiarotti and K. Bostoen. forthcoming. Nouvelle contribution à l'inventaire et la documentation des parlers teke (bantou, B70) en République Démocratique du Congo. *Linguistique et Langues Africaines*.
- Pacchiarotti, S. and K. Bostoen. 2020. The Proto-West-Coastal Bantu velar merger. *Africana Linguistica* 26: 139-195.
- Pacchiarotti, S. and K. Bostoen. 2021. Final vowel loss in Lower Kasai Bantu (DRC) as a contact-induced change. *Journal of Language Contact* 14: 437-474.
- Pacchiarotti, S. and K. Bostoen. 2022. Erratic velars in West-Coastal Bantu: Explaining irregular sound change in Central Africa. *Journal of Historical Linguistics* 12: 1-65.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., & Huelsenbeck, J. P. (2012). MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space. *Systematic Biology*, 61(3), 539–542.
- Rottland, F. 1977. Reflexes of Proto-Bantu Phonemes in Yanzi (B85). *Africana Linguistica* 7: 375-396.
- Seidensticker, D., W. Hubau, D. Verschuren, C. Fortes-Lima, P. de Maret, C.M. Schlebusch and K. Bostoen. 2021. Population collapse in Congo rainforest from 400 CE urges reassessment of the Bantu Expansion. *Science Advances* 7, eabd8352: 1-13.