

Bantu Diachronic Phylogenetics: A case study for the Kikongo Language Cluster

Gilles-Maurice de Schryver, Rebecca Grollemund, Simon Branford and Koen Bostoen

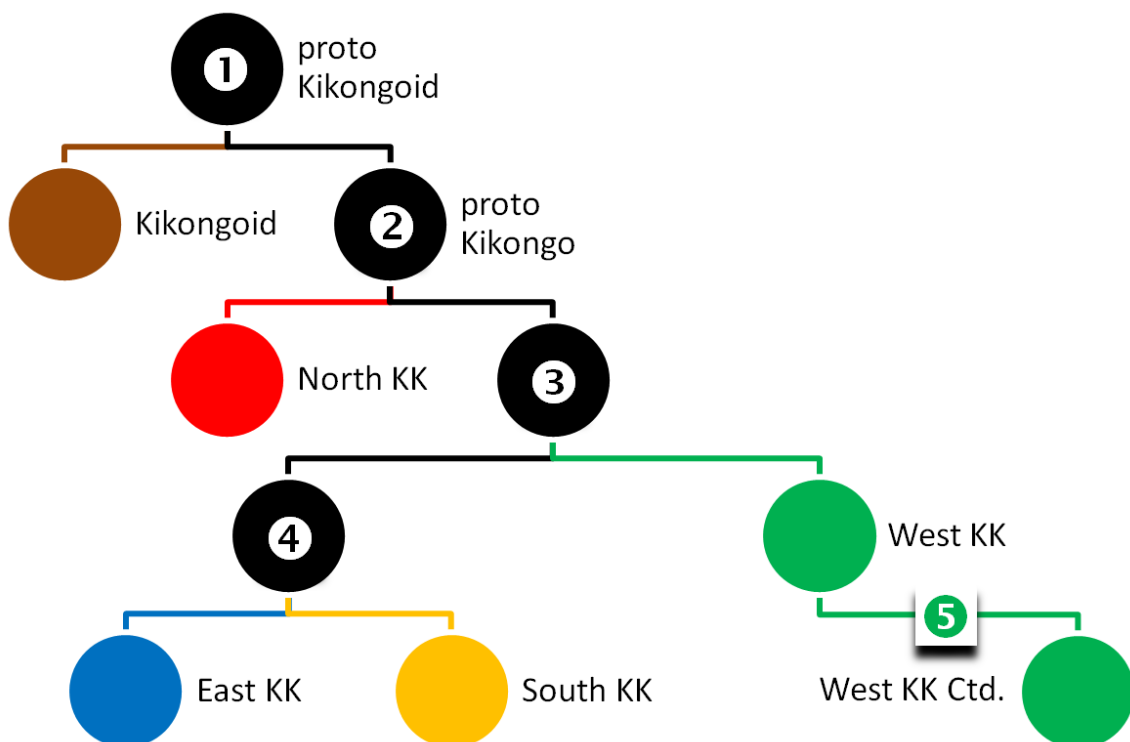
The Lower Congo region in Central Africa occupies a pivotal position in Bantu history. Received knowledge tells us that it lies at the junction of two major Bantu subgroups, the North-West Bantu and West Bantu, not only geographically, but also historically. Several early studies have also located a secondary nucleus of early Bantu expansion in this zone (Heine *et al.* 1977, Vansina 1984). Similarly, a lexical classification of Guthrie's zone C by correspondence analysis has suggested that the coastal Congo area is at the start of a long-term wave-of-advance which resulted in the settlement of the Inner Congo Basin by Bantu speakers (Wotzka & Bostoen 2009). However, it remains very poorly understood how and when Bantu speech communities peopled the Lower Congo region itself.

So as to elucidate the language dynamics in this region, we present a new lexically-driven phylogenetic classification of the 'West-Coastal Bantu languages' (Vansina 1995), consisting of the languages belonging to Guthrie's zone H and groups B40-80 with a special focus on Kikongo varieties (H10). In our dataset we integrated new Kikongo varieties collected during recent fieldwork (2012-2015) as well as all historical Kikongo varieties for which we could find extant data. The actual use of the latter is truly novel in Bantu studies. Our record starts nearly 400 years ago (Cardoso 1624, van Gheel 1652, Brusciotto 1659), and further contains material from about 200 years ago (Descourvières 1772, 1773, 1775, 1776, Tuckey 1818, Koelle 1854), as well as over 15 sources for Kikongo varieties spoken a century ago, and about 50 present-day varieties. We also included a further 50+ surrounding West Bantu languages, both synchronic and diachronic varieties (e.g., de Canneattim 1805), as well as a handful of Grassfields and zone A languages from the North-West.

In order to infer our tree, we used advanced computational phylogenetic methods. We applied a likelihood model of lexical evolution (which makes allowance for different rates of evolution for the studied concepts) and Bayesian inference of phylogeny (using Markov chain Monte Carlo -MCMC) with relaxed clock dating methods, calibrated with the historical Kikongo sources on the one hand, and archaeological evidence on the other. This method allowed us to provide a topology and date estimates for all nodes.

The results regarding the external classification show that the Lower Congo region is an area of later settlement for a subset of West-Coastal Bantu languages, rather than a West-Coastal homeland as claimed in the literature; whereas the results for the internal classification of the Kikongo varieties has revealed the existence of what we want to term the 'Kikongo Language Cluster' (KLC). The KLC is a disparate continuum of closely related Bantu languages spoken in the wider Lower Congo region and constitutes a discrete clade within West-Coastal Bantu. In addition to the H10 languages, this clade fully encompasses the B40 languages spoken as

far away as Gabon, and even includes what we wish to term 'Kikongoid' languages such as Kiyaka (H31), Kisuku (H32), Hungan (H42) and Samba (L12a) spoken in the Kwilu-Kwango region (i.e., some H30 and H40 languages, as well as L12a, but not H20 languages). The core Kikongo (KK) Language Cluster further fragments into four neat subgroups, for which we use cardinal labels: North KK, East KK, South KK and West KK. The Congo and Inkisi Rivers agreeably divide the Kikongo varieties within these sub-groups, with North KK and West KK strictly to the north of the Congo River, East KK strictly to the east of the Inkisi River, and South KK strictly to the south of the Congo River (bar one riparian variety). A schematic representation of the relationship between the sub-groups within the KLC is as follows:



Our multiple diachronic experiments indicate that parent languages of each of these four sub-groups and the sub-groups themselves were already in existence about 1 500 years ago. This time-depth thus considerably predates the emergence of the Kongo kingdom, which oral traditions place in the second half of the 14th century. The state-formation processes that accompanied the rise of the Kongo kingdom have thus no bearing on the existence of those four sub-groups within the KLC, nor is the wide dispersion of the more than 30 core KLC varieties a consequence of the Civil War during the second half of the 17th–early 18th century (Thornton 1983). There is one region where the expanding Kongo kingdom *did* have a noticeable classificatory effect on the internal divisions of the KLC, however, as it even led to the emergence of a contact zone where South KK, East KK and North KK met, creating a convergence zone. We have termed this emerging contact zone, situated north and south of the unnavigable stretches of the Congo River, Central KK.

Finally, our phylogenetic classification of the KLC also allows us to tentatively propose and date the migrations which underlay its emergence; cf. the map below.

References

- Brusciotto, G. 1659. *Regulae quaedam pro difficillimi Congensium idiomatis faciliiori captu ad grammaticae normam redactae*. A F. Hyacintho Brusciotto a Vetralla Concionatore Capuccino Regni Congi Apostolicae Missionis Praefecto. Rome: Sacra Congregatio de Propaganda Fide.
- Cardoso, M. 1624. *Dovtrina Christaã. Composta pelo P. Marcos Iorge da Companhia de IESV Doutor em Theologia. Acrescentada pelo Padre Ignacio Martinz da mesma Companhia Doutor Theologo. De nouo traduzida na lingoa do Reyno de Congo, por ordem do P. Matheus Cardoso Theologo, da Campanhia de IESV*. Lisbon: Geraldo da Vinha.
- de Cannecattim, B. M. 1805. *Collecção de observações grammaticaes sobre a lingua Bunda ou Angolense compostas*. Lisbon: Impressão Regia.
- Descourvières, J. J. 1772. *Dictionnaire françois et congo* (London, manuscript n° 33779, copyist Belgarde).
- . 1773. *Dictionnaire françois et congo* (Besançon, manuscript n° 525, copyist Cuénot).
- . 1775. *Dictionnaire congo et françois* (Besançon, manuscript n° 524, copyist Cuénot).
- . 1776. *Essai d'une grammaire congo, suivant l'accent de Kakongo* (Besançon, manuscript n° 523, copyist Cuénot).
- Heine, B., H. Hoff & R. Vossen. 1977. 'Neuere Ergebnisse zur Territorialgeschichte der Bantu'. In W. J. G. Möhlig, F. Rottland & B. Heine (eds), *Zur Sprachgeschichte und Ethnohistorie in Afrika. Neue Beiträge afrikanistischer Forschungen*, 57–72. Berlin: Dietrich Reimer Verlag.
- Koelle, S. W. 1854. *Polyglotta Africana, or a Comparative Vocabulary of Nearly Three Hundred Words and Phrases in More than One Hundred Distinct African Languages*. London: Church Missionary House.
- Thornton, J. K. 1983. *The Kingdom of Kongo: Civil war and transition, 1641-1718*. Madison, WI: University of Wisconsin Press.
- Tuckey, J. H. 1818. *Narrative of an Expedition to Explore the River Zaire, Usually Called the Congo, in South Africa, in 1816, under the direction of Captain J. K. Tuckey, R. N. To which is added, the journal of Professor Smith; some general observations on the country and its inhabitants; and an appendix: containing the natural history of that part of the kingdom of Congo through which the Zaire flows*. London: John Murray.
- van Gheel, J. 1652. *Vocabularium Latinum, Hispanicum, e Congense. Ad Usum Missionariorû transmittendorû ad Regni Congo Missiones* (manuscript kept in the Biblioteca nazionale centrale di Roma: Fondi minori 1896, Mss. Varia 274).
- Vansina, J. 1984. 'Western Bantu expansion', *Journal of African History* 25(2):129–45.
- . 1995. 'New linguistic evidence and 'the Bantu expansion'', *Journal of African History* 36(2):173–95.
- Wotzka, H.-P. & K. Bostoen. 2009. *Human settlement and Bantu language dispersal in the Inner Congo Basin: A correspondence (re-)analysis of lexicostatistical data*. Paper presented at the 3rd European Conference on African Studies, Leipzig.