

## LINGUISTIQUE HISTORIQUE

Koen Bostoën<sup>1</sup>

Les linguistes et les archéologues développent des points de vue complémentaires sur la culture et le comportement des sociétés africaines du passé. Tandis que la linguistique historique comparée traite habituellement des traces immatérielles du passé dans les langues de l'Afrique contemporaine, l'archéologie met au jour des vestiges matériels de cultures anciennes. Même si les deux disciplines partagent des concepts clés similaires, leurs méthodes, données et cadres interprétatifs diffèrent profondément. Cette contribution vise à expliquer les principes de base de la linguistique historique comparée – telle qu'appliquée aux langues bantu – et à débusquer une série de fausses idées courantes. Sa longueur étant limitée, je ne fournirai pas de références bibliographiques détaillées dans le texte (voir mes précédentes publications pour une bibliographie extensive)<sup>2</sup>. Quelques lectures essentielles pour les non-spécialistes sont listées à la fin du chapitre.

## I. LINGUISTIQUE DIACHRONIQUE BASÉE SUR DES DONNÉES DIACHRONIQUES

Idéalement, la linguistique historique est l'étude de stades historiques distincts de l'évolution d'une langue particulière ou d'une famille de langues. C'est le cas, par exemple, des langues romanes, pour lesquelles on peut reconstruire empiriquement l'évolution du latin en multiples langues filles. En Afrique, l'examen de la variation linguistique à travers le temps, sur la base de données linguistiques diachroniques, n'est pas faisable en raison du manque de documents écrits.

Le cas du kikongo, dont les mentions historiques remontent au début du XVII<sup>e</sup> siècle, est exceptionnel, dépassant même celui du kiswahili dont les textes disponibles les plus anciens ne datent pas au-delà du milieu du XVIII<sup>e</sup> siècle. Pour la plupart des autres langues d'Afrique centrale, les documents écrits n'existent, au mieux, qu'à partir de la fin du XIX<sup>e</sup> siècle. Aujourd'hui encore, il reste par ailleurs des langues non documentées, dont plusieurs sont au bord de l'extinction. En Afrique, la linguistique historique passe donc par l'étude comparée de langues historiquement liées. Cette approche « à rebours » également nommée « linguistique historique comparée », part des langues existantes et essaie

de reconstruire leur évolution à partir des stades originels, en étudiant les variations actuelles entre les langues. Celles-ci peuvent être phonologiques, morphologiques, syntaxiques, sémantiques ou lexicales.

Dans le cas des langues bantu, l'hypothétique langue originelle commune reconstruite sur la base des similarités observées entre des langues connues principalement à partir du XIX<sup>e</sup> siècle est appelée « proto-bantu ». Cette protolangue est considérée comme le meilleur reflet possible d'une langue ancestrale qui a pu être parlée il y a 4 000 à 5 000 ans, dans la zone à partir de laquelle les langues ont commencé leur expansion à travers l'Afrique centrale et au-delà. Les linguistes spécialistes du bantu s'accordent pour situer ce foyer originel dans la région des Grassfields au Cameroun, non loin de la frontière avec le Nigeria. Cette zone présente la diversité linguistique la plus élevée (ce qui signifie que les langues apparentées ont eu suffisamment de temps pour diverger localement) et elle est proche de la zone où sont parlées les langues benue-congo, auxquelles sont apparentées les langues bantu.

## II. CLASSIFICATIONS RÉFÉRENTIELLES VS HISTORIQUES OU GÉNÉALOGIQUES

La classification des langues bantu la plus connue est sans nul doute celle de Malcolm Guthrie. En 1948, Guthrie a subdivisé les langues bantu en 16 zones distinctes, étiquetées A, B, C, D, E, F, G, H, K, L, M, N, P, R, S et T, qu'il a réduites à 15 en 1971 en fusionnant les deux dernières. Chaque zone est ensuite subdivisée en groupes linguistiques indexés à un nombre décimal et dans lesquels une unité est assignée à chaque langue. Les lettres minuscules accolées à certaines unités renvoient à des dialectes de la même langue, par exemple ciluba (L31a) et luluwa (L31b). Contrairement à ce que l'on croit souvent, la classification de Guthrie n'est pas strictement référentielle et n'a jamais été conçue comme historique : Guthrie ne s'appuyait pas sur la « méthode comparative » (qui est l'approche centrale de la linguistique historique comparée) ou sur les « innovations partagées », principe de base d'identification de sous-groupes historiques. Les innovations partagées sont des changements lexicaux, phonologiques ou grammaticaux survenus une seule fois dans une langue originelle et qui ont été transmis à ses langues filles, ce qui en fait des indicateurs de l'affiliation plus étroite du lien entre langues. En attribuant un code alphanu-

1 BantUGent - Centre d'Études bantoues de l'UGent, Département de Langues et Cultures, Université de Gand, Belgique.

2 <http://research.flw.ugent.be/en/koen.bostoën>

mérique unique à chaque langue, Guthrie voulait faciliter les comparaisons entre les centaines de langues bantu connues à l'époque.

En dépit de sa valeur historique ou généalogique limitée, la classification de Guthrie reste un outil de référence utile. Chacune des quelque 900 variétés de langues bantu documentées peut être approximativement localisée dans l'espace grâce à son code unique. C'est précisément pour cela que Jouni Maho a actualisé la liste de Guthrie, en lui ajoutant de nouvelles langues, tout en restant le plus fidèle possible à son approche originale. D'autres chercheurs ont proposé des réaménagements pour des raisons historiques. Seule l'une de ces modifications a été assez largement adoptée par la communauté des spécialistes du bantu, qui concerne la zone J et a été proposée par l'ancien département de Linguistique du Musée royal de l'Afrique centrale à Tervuren.

En tant que famille linguistique, le bantu a été reconnu depuis Bleek (1851). Son berceau se situe dans la région où se rencontrent le bantu dit « étroit » (« *Narrow Bantu* »), à savoir les langues classées comme bantu par Guthrie, et le bantu dit « élargi » (« *Wide Bantu* »), à savoir leurs parents Benue-Congo les plus proches aussi appelés « bantoïdes ». Le petit sous-groupe « Mbam-Bubi », qui se compose de plusieurs langues parlées dans la région du Mbam au centre du Cameroun ainsi que du bubi de l'île de Bioko, est la jointure généalogique entre le bantu étroit et le bantu élargi. La famille bantu au sens strict se subdivise en cinq autres sous-groupes majeurs: « Nord-Ouest », « Centre-Ouest » (aussi connu comme « Zaïre du Nord » ou « Congo »), « Ouest-Ouest » (aussi connu comme « Côte-Ouest », « Sud-Ouest ») et « Est ». Cette compréhension élémentaire de la généalogie bantu se base avant tout sur les études quantitatives du vocabulaire dit « de base », comme la lexicostatistique et la phylogénétique. Les approches qualitatives s'appuyant sur les données phonologiques et/ou grammaticales sont moins compatibles avec le modèle arborescent de diversification linguistique et soulignent que la convergence due au contact linguistique avait également un impact important sur la spéciation des langues bantoues.

### III. LA LANGUE COMME SOURCE HISTORIQUE

Notre connaissance des phénomènes environnementaux, sociaux, culturels et historiques qui sous-tendent les changements linguistiques est, dans l'ensemble, très restreinte pour l'Afrique. Les langues du continent doivent le plus souvent « parler pour elles-mêmes ». L'étude des langues en tant que telle est devenue une approche importante de reconstruction de l'histoire, à laquelle travaillent non seulement les lin-

guistes mais aussi les historiens et les archéologues. Fondée sur le principe de base selon lequel le vocabulaire commun à des communautés linguistiques<sup>4</sup> reflète une histoire partagée, l'étude de termes culturels répandus procure habituellement des informations intéressantes sur le mode de vie des sociétés du passé. Cette sous-discipline est aussi connue comme la méthode « des mots et des choses » (voir Ricquier, ce volume, pp. 261-263) ou la « paléontologie linguistique ». Pour les archéologues, les données linguistiques sont particulièrement utiles comme sources historiques indirectes pour les aspects immatériels des cultures humaines ou pour les traces matérielles qui se conservent mal. Des mots proches dotés de sens similaires et communs à de nombreuses langues peuvent avoir été hérités d'une langue originelle commune et s'être diffusés *via* la dispersion de ses langues filles. Ils peuvent également avoir été adoptés par contact et s'être diffusés à travers les langues comme mots d'emprunt.

Pour faire la distinction entre vocabulaire hérité et emprunté, les linguistes se basent sur le principe des correspondances phonétiques régulières. Ce sont des similarités phonologiques entre langues qui ne peuvent pas être le résultat d'accidents historiques, car elles sont récurrentes, systématiques et sans exceptions inexplicables. Alors que les termes hérités et largement distribués peuvent être reconstruits dans la protolangue hypothétique, en s'appuyant sur les changements phonétiques, ce n'est pas le cas des mots empruntés. Plusieurs langues bantu des Grands Lacs ont, par exemple, des doublets lexicaux pour signifier « calebasse » et « bouteille de verre ». Ce sont deux mots historiquement reliés, mais l'un d'eux a été acquis par transmission intergénérationnelle régulière à partir d'une langue ancestrale commune, tandis que l'autre a été obtenu à partir du swahili véhiculaire selon un processus de diffusion par contact. Le terme hérité pour « calebasse » est beaucoup plus hétérogène phonologiquement : par exemple « *cuβa* » pour le sukuma, « *nsòhá* » en nyamwezi, « *énsúwà* » en ganda, « *nshùhá* » en shi. Ces mots ont été sujets aux changements phonétiques réguliers que leurs langues ont subis à partir du proto-bantu pour lequel « *\*-cópà* » (« calebasse ») avait été reconstruit. Ce n'est pas le cas pour le terme signifiant « bouteille de verre » qu'ils ont récemment emprunté au swahili sous la forme de mots d'emprunt beaucoup plus proches : « *cupá* » en sukuma, « *cupa* » en nyamwezi, « *ccúpà* » en ganda, « *icúpà* » en shi. En swahili même, le mot « *chupa* » signifie à la fois « calebasse » et « bouteille de verre ». Quand ces derniers types de contenants furent introduits le long de

4 On définit ici une communauté linguistique comme un groupe de personnes qui estiment parler une même langue.

la côte est-africaine, les locuteurs swahili les nommèrent d'après leurs récipients traditionnels, utilisant pour « calebasse » le terme qu'ils avaient hérités du proto-bantu. Les commerçants swahili-phones engagés dans les échanges à longue distance introduisirent par la suite ce nouvel élément de culture matérielle et les mots swahili les désignant dans plusieurs communautés est-africaines, nombre d'entre elles disposant déjà d'un terme bantu régulièrement hérité pour « calebasse ».

À la différence des archéologues, les linguistes ne disposent pas d'une méthode standard et universellement acceptée en matière de datation absolue des changements linguistiques, même s'ils peuvent utiliser des données archéologiques comme points de calibration pour générer des phylogénies datées comme l'ont récemment essayé Grollemond *et al.* (2015). En l'absence de données linguistiques diachroniques et de mise en relation des données linguistiques et archéologiques, les linguistes sont obligés de s'en tenir à une datation relative. Pour ce faire, ils s'appuient sur une série de principes empruntés à l'archéologie : stratigraphie, distribution géographique et sériation.

Les linguistes recourent au concept de stratigraphie pour désimbriquer les couches successives de la formation d'une langue. La grammaire et le lexique d'une langue sont transmis dans la durée et transformés par perte d'éléments anciens et incorporation d'éléments neufs. Ils accumulent des couches de formation qui ne se superposent jamais de manière nette. Contrairement aux strates archéologiques, les couches linguistiques ne suivent pas la loi de la superposition. Il se produit une sorte de contamination stratigraphique permanente. Il revient au chercheur en linguistique historique d'ordonner les données actuelles en strates successives. Les mots pour « calebasse » et « bouteille de verre » de l'exemple précédent appartiennent clairement à deux couches distinctes d'histoire linguistique.

La géographie linguistique ou géolinguistique peut contribuer à la datation relative des couches linguistiques. Cette méthode traite de la distribution géographique des traits linguistiques. Elle est utilisée pour cartographier les routes de diffusion des mots empruntés et pour déterminer la direction de l'emprunt, mais aussi comme instrument de chronologie relative. Les « isoglosses linguistiques » sont les équivalents des horizons stylistiques en archéologie. Ils signalent la distribution géographique d'un trait linguistique donné, partagé par plusieurs langues. Par exemple, les mots pour « oiseaux » apparentés à « *njila* » (en kimbundu) n'apparaissent que dans un groupe géographiquement restreint de langues bantu parlées dans la partie sud-ouest

du domaine bantu, tandis que les mots apparentés au mot « *nuni* » kikongo signifiant « oiseau », « *nuni* », se rencontrent partout dans cet espace. On interprète une telle distribution en fonction du temps : le mot kikongo est une survivance commune remontant au proto-bantu, alors que le terme kimbundu est une innovation partagée plus récente. On procède à l'interprétation des isoglosses en termes de chronologie relative, selon certaines normes aréales qui ne sont pas des règles strictes mais plutôt des principes herméneutiques, selon lesquels, par exemple, la forme la plus ancienne est la plus dispersée, ce qui est préférentiellement attesté dans les zones les plus périphériques, tandis que la forme la plus récente apparaît dans un groupe de langues adjacentes, qui peut être vaste, mais pas aussi dispersé que pour la forme plus ancienne. Une interprétation historiquement judicieuse des isoglosses requiert une connaissance de base de la classification interne de la famille de langues. La profondeur temporelle relative ne dépend pas tant du nombre de langues dans lesquelles on constate l'occurrence d'un trait que de la distribution de ce dernier dans des sous-groupes historiquement distincts. Dès lors, un terme rare mais dispersé à travers les langues bantu du nord-ouest et de l'ouest sera tenu pour antérieur à un synonyme dont la présence est dense mais restreinte aux langues bantu de l'est.

Un concept archéologique de base également utilisé par la linguistique historique est la sériation. Les linguistes s'en servent habituellement pour l'ordonnement séquentiel des changements phonétiques. Toute langue est sujette à des changements phonétiques que l'on peut qualifier de réguliers, dans la mesure où ils affectent tous les mots partageant un environnement phonologique donné. Le séquençage chronologique des changements phonétiques est principalement utilisé dans la classification historique des langues suivant le principe des innovations partagées. Si des langues étroitement apparentées partagent un changement historique (qu'il soit lexical, phonologique ou grammatical), il y a une forte chance que cette innovation ne soit survenue qu'une fois, et donc au sein de leur langue parente la plus récente – bien que la convergence de changements indépendants ne puisse jamais être totalement exclue. Une fois qu'on a une idée de la classification interne d'un groupe de langues et de la chronologie relative des modifications phonétiques, la sériation peut aussi servir à la datation des termes d'emprunt. Plus les mots étrangers ont été empruntés tôt, plus ils auront de changements phonétiques en commun avec les mots régulièrement hérités, et meilleure sera leur intégration phonologique, rendant difficile leur identification en tant que vocabulaire d'emprunt.

## CONCLUSIONS

L'interaction entre archéologie et linguistique africanistes a été sévèrement critiquée par le passé en raison, entre autres, de l'absence d'évaluation critique des méthodes et concepts sous-jacents. Même si cette appréciation n'est sans doute pas imméritée, elle ne devrait pas empêcher la collaboration interdisciplinaire. Aucune discipline n'est capable de résoudre à elle seule les énigmes complexes de l'histoire africaine. Un bon travail d'équipe linguistico-archéologique nécessite tout d'abord une bonne compréhension mutuelle des concepts, méthodes et données de chacun, ce à quoi j'ai essayé de contribuer dans ce chapitre. Un second enjeu fondamental réside dans l'importance d'une collaboration directe entre chercheurs de différentes disciplines qui maîtrisent parfaitement leurs propres corpus de données et sont à même de prononcer des avis judicieux quant à leur signification historique, plutôt que de laisser ce travail à des chercheurs qui ne maîtrisent qu'une méthode, voire n'en maîtrisent aucune. Enfin, il est crucial qu'archéologues et linguistes bénéficient mutuellement de leurs avantages spécifiques, par exemple la datation absolue dans le cas de l'archéologie, ou la possibilité de reconstruire un vocabulaire en s'appuyant sur les dimensions immatérielles ou mal préservées de la vie humaine pour la linguistique historique.

## BIBLIOGRAPHIE

- Bastin, Y. & Piron, P. 1999. « Classifications lexicostatistiques : bantou, bantou et bantoïde. De l'intérêt des "groupes flottants" ». In J.M. Hombert & L.M. Hyman (éd.), *Bantu historical linguistics: theoretical and empirical perspectives*. Stanford: CSLI Publications, pp. 149-164.
- Bastin, Y., Coupez, A. & Mann, M. 1999. *Continuity and Divergence in the Bantu Languages : Perspectives from a Lexicostatistic Study*. Tervuren : MRAC (coll. « Annales de Sciences humaines, série in-8° », n° 162).
- Bleek, W. 1851. *De nominum generibus linguarum Africae Australiae*. Bonnae : Formis Caroli Georgii.
- Blench, R. 2006. *Archaeology, Language and the African Past*. Lanham : Altamira Press.
- Bostoen, K., Clist, B., Doumenge, C., Grollemund, R., Hombert, J.-M., Koni Muluwa, J. & Maley, J. 2015. « Middle to Late Holocene Paleoclimatic Change and the Early Bantu Expansion in the Rain Forests of Western Central Africa ». *Current Anthropology* 56 (3) : 354-384.
- Dimmendaal, G. 2011. *Historical Linguistics and the Comparative Study of African Languages*. Amsterdam/Philadelphia : John Benjamins.
- Grollemund, R., Branford, S., Bostoen, K., Meade, A., Venditti, C. & Pagel, M. 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*, <http://dx.doi.org/10.1073/pnas.1503793112>
- Heine, B. & Nurse, D. (éd.). 2000. *African Languages : An Introduction*. Cambridge/New York : Cambridge University Press.
- Heine, B. & Nurse, D. (éd.). 2004. *Les Langues africaines*. Paris : Karthala.
- Hombert, J.-M. & Hyman, L.M. (éd.). 1999. *Bantu Historical Linguistics : Theoretical and empirical perspectives*. Stanford, CA : CSLI.
- Maho, J.-F. 2009. *NUGL Online : The online version of the New Updated Guthrie List, a referential classification of the Bantu languages*. (dossier en ligne : <http://goto.glocalnet.net/mahopapers/nuglonline.pdf>, version du 4 juin 2009).
- Nurse, D. & Philippson, G. (éd.). 2003. *The Bantu languages*. London/New York : Routledge (coll. « Routledge Language Family » 4).
- Schadeberg, T.C. 2003. « Historical linguistics ». In D. Nurse & G. Philippson (eds.), *The Bantu Languages*. London/New York : Routledge, pp. 143-163.
- Vansina, J. 1995. « New Linguistic Evidence and the Bantu Expansion ». *Journal of African History* 36 (2) : 173-195.