

Description du projet *Création d'une base empirique pour la typologie aréale des langues bantu du nord-ouest (BantuTyp).*

(extrait de la demande)

budget : 146 000€

Introduction

Le but principal du projet BantuTyp est de créer une analyse grammaticale de trois langues bantu parlées dans la même zone géographique (le nord-ouest de l'aire bantu), qui sont – évidemment – génétiquement apparentées, mais qui diffèrent l'une de l'autre au niveau typologique.

Ces analyses grammaticales doivent en premier lieu servir de base empirique pour les recherches en typologie et en linguistique aréale, où la question principale est celle de savoir quels traits morphosyntaxiques et phonologiques se répandent indépendamment et quels traits aréaux sont le résultat d'une coévolution suite à l'expansion initiale d'un nombre de traits aréaux plus réduit. Par exemple, parmi les traits aréaux qu'on retrouve couramment dans les langues bantu du nord-ouest, on peut noter l'accent à l'initiale du thème réalisé en termes de longueur consonantique, les contraintes de maximalité sur les thèmes et l'existence de plosives labio-vélaires. La coexistence de ces traits peut être due au fait qu'ils se répandent facilement dans une situation de contact, mais elle peut également être le résultat d'une coévolution dans les langues qui ont acquis un accent initial, qui aurait ensuite déclenché l'apparition des labiovélares (d'abord) à l'initiale du thème, ainsi que l'érosion du matériel segmental en finale de thème.

Toutefois, la pertinence du projet proposé va bien au-delà de cette question de recherche précise. Il nous manque des analyses grammaticales accessibles, englobantes et fiables – c.-à-d. reproductibles – des langues du monde en général et des langues de l'Afrique en particulier. Cette situation affaiblit la base empirique de toutes les approches en linguistique qui se veulent typologiques, au sens le plus large du terme. Ainsi, pour la famille bantu le nombre de grammaires de référence se compte sur une main et la seule description véritablement englobante, celle du Mongo (Hulstaert 1961, 1965, 1966), est peu accessible aux non-bantuisants. Pourtant, la famille bantu est la plus grande du monde (au niveau de "genus"), et probablement la famille linguistique la mieux étudiée de l'Afrique noire.

Si le manque d'analyses grammaticales englobantes des langues du monde est problématique, le manque de fiabilité (falsifiability) des analyses existantes l'est

davantage. Faute de documentation accompagnante, l'utilisateur des analyses grammaticales doit se contenter généralement des exemples fournis par le grammairien pour soutenir ses analyses, des exemples qui sont en plus réduits à l'écriture, de sorte que certaines informations pertinentes peuvent être omises (ton, accent, longueur, etc).

Les grandes familles linguistiques ont l'avantage unique qu'on peut y étudier les évolutions divergentes de systèmes grammaticaux qui à l'origine étaient à peu près identiques, pour ainsi mieux comprendre les corrélations entre traits typologiques et le rôle de phénomènes de contact (voir aussi Dunn et al, 2011).

Nous sélectionnerons deux langues sur la base de leurs propriétés typologiques. Une langue qui a mieux préservé la structure grammaticale reconstruite en proto-bantu et une langue qui a subi la majorité des changements typiquement associés avec les langues du nord-ouest. Un candidat de langue du premier type est le bapuku, langue côtière du groupe A30 parlée au sud-ouest du Cameroun, pour laquelle le porteur principal a fait une esquisse grammaticale et lexicale en 2012. Le groupe appelé A90 est un bon candidat pour fournir une langue du deuxième type. Les langues de ce groupe, à savoir le pol, le kwakum et le kaka, ne sont quasiment pas décrites.

La troisième langue qui sera décrite au sein du projet BantuTyp est le myènè, langue pour laquelle nous avons déjà établi une documentation importante au sein d'un projet financé par le Endangered Languages Documentation Fund. Voir le site du projet pour plus de détails : <http://llacan.vjf.cnrs.fr/myene/>. Cette langue sera décrite principalement par Odette Ambouroué et Mark Van de Velde.

Livrables

- Une thèse de doctorat qui consistera d'une analyse grammaticale, un dictionnaire de 3000 entrées et une documentation d'une langue bantu du groupe A90.
- Une analyse grammaticale de base, un dictionnaire et une documentation de la langue bapuku (potentiellement en synergie avec un projet de documentation ELDP dont la demande est en cours, projet qui sera basé à la SOAS à Londres).
- Les bases (publiées en ligne) d'une grammaire de référence de l'ensemble dialectal myènè, qui a l'ambition de devenir l'analyse grammaticale la plus englobante de la famille bantu.

- Un nombre d'articles sur des phénomènes phonologiques et grammaticaux remarquables du point de vue de la linguistique aréale ou de la typologie linguistique.

Questions de recherche

Il y a un nombre de phénomènes grammaticaux qui intéressent particulièrement les porteurs de la présente demande et qui sont pertinents pour les projets de recherche dans lesquels ils sont impliqués, mais qui ne peuvent pas être étudiés dans une langue en l'absence d'une description intégrée de la phonologie, tonologie, morphologie et syntaxe. Notre but est donc de créer des analyses grammaticales englobantes qui font particulièrement attention aux questions de recherche suivantes, toutes inspirées par une question fondamentale: *quels traits phonologiques et grammaticaux sont le résultat d'une évolution interne à la famille bantu et lesquels sont dus à l'introduction de traits aréaux?*

- Y a-t-il un accent de mot à l'initiale du thème, sur la syllabe pénultième du mot, ou les deux? Est-ce que le corrélat phonétique de l'éventuel accent initial est la longueur consonantique? Si oui, quelle est la différence de longueur moyenne entre les consonnes accentuées et les mêmes consonnes en position non accentuée? Y a-t-il d'autres corrélats phonétiques que la longueur consonantique? Est-ce que la langue a des phonèmes segmentaux dont la présence peut être due à l'accent initial? Y a-t-il une distribution phonotactique asymétrique des segments consonantiques et, si oui, cette distribution est-elle semblable à celle des autres langues de la région qui ont ce phénomène? Quel est le corrélat phonétique d'un éventuel accent pénultième et quel est le domaine exact de cet accent? Est-ce que le domaine d'accentuation est plus grand en débit rapide qu'en débit lent? Est-ce qu'une éventuelle réduction de la voyelle finale crée l'impression d'une voyelle préfinale allongée? Quels règles segmentales et tonales réfèrent à l'accent? Y a-t-il des contraintes de maximalité sur le nombre de syllabes par thème? Si oui, est-ce que ces contraintes sont "productives"? Y a-t-il, par exemple, des règles morphologiques qui réfèrent à la structure syllabique du mot?

- Y a-t-il de l'harmonie vocalique dans la langue? Si oui, Quels sont les processus d'harmonie vocalique ? Quels sont les traits phonologiques en jeu? Est-ce le trait [ATR] ou le trait [RTR]? Quels sont les morphèmes ciblés par l'harmonie vocalique? Est-ce que l'harmonie vocalique est symétrique par rapport à la racine?
- Est-ce que la morphologie verbale dérivationnelle est restreinte par rapport à celle du proto-bantu? Si oui, quelles en sont les conséquences au niveau de la syntaxe de la phrase? Est-ce que les catégories TAM sont exprimées de façon synthétique ou périphrastique? Est-ce que la langue utilise les mêmes stratégies analytiques pour exprimer les notions dites adverbiales? Quelles notions sont grammaticalisées dans le sens que leur expression soit obligatoire? Si on peut démontrer des cycles de grammaticalisation dans le système TAM, est-ce que le système grammatical reste inchangé d'un point de vue notionnel?
- Peut-on trouver des qualifiants qui sont exprimés de la même façon que le possédé dans le syntagme génital? Si oui, y a-t-il différentes constructions de ce genre? Y a-t-il d'autres parallèles entre les constructions qualificatives et les constructions génitales? Y a-t-il des cas de "head marking" de certaines relations grammaticales au sein du syntagme nominal ? En d'autres termes, est-ce que les noms prennent une autre forme s'ils sont modifiés par un démonstratif, un pronom possessif ou une phrase relative ? Si oui, y a-t-il des indications sur l'origine du marqueur sur le nom ? Est-ce que l'ordre des mots est fixe au sein du syntagme nominal. Si non, est-ce que les différentes sérialisations ont des valeurs pragmatiques différentes ?

Synergie avec les projets scientifiques existants

Plutôt que de fédérer un grand nombre de chercheurs de différents horizons au sein d'un projet et de contribuer ainsi à l'éparpillement des forces des chercheurs, le projet BantuTyp constituera une petite équipe pluridisciplinaire (chercheurs en phonologie de laboratoire et en linguistique de terrain), dont les recherches se dérouleront en synergie avec de nombreuses projets de recherche en cours. L'aspect fédérateur du projet BantuTyp se situe donc dans son interaction avec et sa pertinence pour un bon nombre d'autres projets en cours.

Les trois dictionnaires contribueront au projet RefLex (projet ANR, responsable Guillaume Segerer, <http://reflex.cnrs.fr/>), qui a pour but, parmi d'autres, de mieux exploiter les données lexicales sur les langues africaines pour des buts de classification et reconstruction.

L'analyse grammaticale des variétés myènè (bantu B10) s'appuiera fortement sur l'énorme documentation rassemblée lors du projet ELDP Comparative Documentation of the Myene language cluster (<http://llacan.vjf.cnrs.fr/myene/>). De son côté, BantuTyp contribuera aux résultats de ce projet en créant des transcriptions et annotations supplémentaires, qui seront toutes rendues publiques et librement accessibles par le biais de l'archive ELAR.

Le projet BantuTyp doit surtout renforcer les bases empiriques du projet Areal phenomena in northern sub-Saharan Africa (projet LC2 de l'Axe 3 du LabEx EFL), qui vise à tester l'hypothèse selon laquelle il existe un Sprachbund en Afrique centrale.

Finalement, le projet BantuTyp s'articulera avec le projet PPC9 Distinctive features in the brain, the mouth, and the ear. (Responsable R. Ridouane) du LABEX EFL. Celui-ci est consacré à l'étude des traits phonologiques. BantuTyp permettra de mener une investigation des traits harmoniques [ATR] et [RTR].

Méthodologie

Les analyses grammaticales produites au sein du projet BantuTyp seront basées sur des données issues aussi bien de l'observation du langage naturel que de l'expérimentation (enquêtes grammaticales et photos et vidéos d'élicitation). Toutes les données primaires récoltées sur le terrain seront traitées selon les normes récemment développées au sein de la discipline de la linguistique documentaire. Ce qui implique, entre autres choses, une quête constante de la meilleure qualité d'enregistrements ; la création de métadonnées uniformisées ; l'usage exclusif de formats de fichiers non-propriétaires ; le respect des règles éthiques lors de la récolte et de la dissémination de la documentation ; et la publication en ligne et l'archivage sur le long terme de la documentation entière.

Ces meilleures pratiques, trop rarement respectées en linguistique de terrain, doivent permettre à tous les chercheurs intéressés de vérifier l'ensemble de nos analyses et d'exploiter nos données pour leurs propres recherches.

Une collaboration intensive entre le doctorant et les autres membres du projet doit permettre au premier de vite dépasser le stade de l'analyse phonologique et morphologique de base. En abordant collectivement cette étape, très chronophage

pour quelqu'un sans expérience, nous permettront au doctorant de passer à des questions plus complexes et intéressantes.

Contexte / Etat de l'art

- Il y a un intérêt renouvelé pour l'analyse grammaticale des langues du monde et une sensibilisation à l'adoption de meilleures pratiques pour la collecte, la sauvegarde et le partage des données primaires. Cet intérêt s'explique par l'essor de la typologie, dont le succès dépend de la disponibilité de grammaires englobantes et fiables. Il est également inspiré par un sens d'urgence face à la disparition rapide de la diversité linguistique, ce qui est en train de détruire une expérience naturelle irremplaçable.
- Notre connaissance des langues des groupes A90 et A30 est très restreinte. Pour le groupe A90, nous ne disposons d'aucune description grammaticale. La langue la mieux étudiée dans ce groupe est le kako, sur laquelle un membre de la SIL, Urs Ernst, a rédigé un nombre d'études ponctuelles, la majorité non publiées. Pour la langue kwakum une esquisse grammaticale figure dans une thèse en ethnolinguistique (Belliard 2006). Le polri a fait l'objet d'un nombre d'études non-publiées d'un jeune chercheur camerounais, Abraham Wega Simeu. La seule langue du groupe A30 qui ait été décrite est le bubi. Les descriptions du bubi sont toutefois datées et de qualité douteuse. En plus, l'inclusion du bubi dans le groupe A30 est probablement erronée d'un point de vue généalogique. La seule source sur la grammaire du bapuku est un mémoire sur la morphologie nominale (Van Hille 1989).
- L'existence d'une aire linguistique dans le nord de l'Afrique subsaharienne est reconnue depuis longtemps (Westermann 1911, Greenberg 1959). Deux publications récentes ont renouvelé l'intérêt pour ce phénomène. Güldemann (2008) étudie la distribution géographique de six phénomènes grammaticaux et démontre que celle-ci est cohérente. Clements & Rialland (2008) identifient six zones phonologiques en Afrique, dont la zone Soudanique correspond dans ces grandes lignes à l'aire définie par les critères de Güldemann. Les critères déjà proposés doivent être raffinés et des critères additionnels ajoutés afin de pouvoir évaluer le bien fondé de l'hypothèse d'un sprachbund. De même, toute hypothèse sur les aires linguistiques doit prendre en compte les découvertes sur l'histoire du peuplement de la région faites par les autres disciplines, telles que l'archéologie, la génétique des

populations et la géographie (Bickel & Nichols 2006). De même, quand on suppose que la distribution d'un phénomène linguistique est due principalement à la diffusion, un scénario de transfert réaliste doit être proposé, dans l'esprit de Matras (2009). Tout ceci est impossible sans une base empirique plus solide. La plupart des langues dans cette zone sont à peine décrites et beaucoup sont en voie de disparition.

Bibliographie

- Ambourou, Odette (2007). *Eléments de description de l'orungu, langue bantou du Gabon (B11b)*. Thèse de doctorat. Université Libre de Bruxelles.
- Belliard, François, (2006). *Instruments, chants et performances musicales chez les kwakum de l'arrondissement de Doume (Est-Cameroun)*. Etude ethnolinguistique de la conception musicale d'une population de langue bantou A91. Thèse de doctorat. Paris 7.
- Bickel, Balthasar & Johanna Nichols. (2006). *Oceania, the Pacific Rim, and the theory of linguistic areas*. *Berkeley Linguistics Society* 32. [PDF available at <http://www.uni-leipzig.de/~bickel/research/papers>]
- Dunn, Michael, Simon J. Greenhill, Stephen C. Levinson & Russel D. Gray (2011). *Evolved Structure of Language Shows Lineage Specific Trends in Word-Order Universals*. In: *Nature* 473: 79-82.
- Greenberg, Joseph H. (1959). *Africa as a linguistic area*. In: William R. Bascom & Melville J. Herskovits (eds.), *Continuity and change in African cultures*, 15-27. Chicago: University of Chicago Press.
- Güldemann, Tom (2008). *The Macro-Sudan belt: towards identifying a linguistic area in northern sub-Saharan Africa*. In: Heine, Bernd & Derek Nurse (eds) *A Linguistic Geography of Africa*, 151-185. Cambridge: Cambridge University Press.
- Hulstaert, Gustaaf. (1961). *Grammaire du ɩmɔ́ngɔ́*. Première partie. Phonologie. Tervuren: Royal Museum for Central Africa.
- Hulstaert, Gustaaf. (1965). *Grammaire du ɩmɔ́ngɔ́*. Deuxième partie. Morphologie. Tervuren: Royal Museum for Central Africa.
- Hulstaert, Gustaaf. (1966). *Grammaire du ɩmɔ́ngɔ́*. Troisième partie. Syntaxe. Tervuren: Royal Museum for Central Africa.
- Hyman, Larry M. (1989). *Accent in Bantu: an appraisal*. In: *Studies in the Linguistic Sciences* 19 (2): 115-134.

- Hyman, Larry M. (2004). How to become a Kwa verb. In: *Journal of West African Languages* 30 (2): 69–88.
- Matras, Yaron (2009). *Language contact*. Cambridge: Cambridge University Press.
- Van Hille, Marion (1989). *Elements de description du syntagme nominal en puku, langue bantoue de zone A*. Mémoire de DEA. Université Libre de Bruxelles.
- Westermann, Dietrich. (1911). *Die Sudansprachen: eine sprachvergleichende Studie*. Abhandlungen des Hamburgischen Kolonialinstituts, 3. Hamburg: L. Friederichsen.