

**Accueil pour activités de recherche
août 2007 - août 2008**

Bilan

Je suis ingénieur linguiste (IE), titulaire au LLACAN depuis septembre 2003. Cependant, je fréquente cette équipe depuis 1995, lorsque je me suis inscrit en DEA de Sciences du Langage sous la direction de France Cloarec-Heiss, alors directrice de l'unité. Depuis plus de 10 ans, j'ai participé à la vie du laboratoire sous presque tous les statuts possibles : d'abord Contrat Emploi Solidarité, puis vacataire, puis CDD, ingénieur d'études stagiaire et enfin ingénieur d'études titulaire. Parallèlement, avant d'être définitivement recruté, j'ai été membre du laboratoire en tant que doctorant puis post-doctorant. Mon implication dans les activités du LLACAN s'est donc construite depuis le début suivant deux axes : des activités techniques et une collaboration scientifique.

Ma formation en linguistique a suivi la voie universitaire standard. En revanche, mes compétences techniques, en particulier en informatique, sont le résultat d'une formation autodidacte ininterrompue depuis une quinzaine d'années.

Bien que je sois resté en contact permanent avec la recherche, mon statut d'ingénieur et mes fonctions au sein du laboratoire ne me permettaient qu'une activité limitée dans ce domaine. C'est pourquoi, en juillet 2005, j'ai sollicité un détachement de deux ans pour activités de recherche au LLACAN. Après des péripéties qu'il n'est pas nécessaire de rappeler ici, ce détachement m'a finalement été accordé en août 2007 pour une durée d'un an, soit jusqu'au 31 août 2008.

Depuis ma thèse (description de la langue bijogo de Guinée Bissau) je m'intéresse aux langues Atlantiques, qui constituent la branche extrême-occidentale du phylum Niger-Congo. Ces langues posent des problèmes particuliers pour le comparatisme, la reconstruction et la classification. Le projet proposé pour l'obtention du détachement est donc naturellement connecté à ces questions. Après un bref rappel de ce projet, je vais donc présenter un premier bilan des travaux effectués dans le cadre de ce détachement.

Le projet initial

Le programme de recherches sur lequel le détachement m'a été accordé porte sur l'établissement des liens généalogiques entre une quinzaine de langues d'Afrique de l'Ouest. Ce groupe de langues, connu sous le nom de BAK dans la littérature (Sapir 1971) occupe une position privilégiée au sein de la branche Atlantique du phylum Niger-Congo. En effet, il est à la fois au centre du domaine et linguistiquement très

morcelé quoique relativement homogène, ce qui fait de la reconstruction du proto-BAK une étape indispensable vers la reconnaissance d'un encore très hypothétique proto-Atlantique.

La division du groupe BAK en trois branches est admise depuis Sapir (1971). Ces branches sont : 1. l'ensemble Joola, dont on identifie une dizaine de variétés ; 2. l'ensemble Manjaku, qui compte de deux à cinq parlers distincts ; 3. le balante, dont les deux dialectes connus semblent assez divergents. Toutefois, dès l'élaboration du projet, il est apparu que la langue *bayot*, jusqu'ici rangée dans l'ensemble Joola, devait en être séparée pour former une quatrième branche.

Le projet avait donc pour objectif d'examiner en détail chacune des branches, de proposer des reconstructions lexicales et grammaticales obtenues par la méthode comprative classique, et enfin de réunir ces reconstructions pour éventuellement élaborer la reconstruction proto-BAK proprement dite.

Le bilan en avril 2008

L'originalité du projet initial était de s'appuyer sur de solides compétences en informatique, avec des approches novatrices dans le traitement automatique des corpus lexicaux. L'expertise accumulée dans l'élaboration et la gestion de bases de données linguistiques m'a en effet fait envisager des développements importants dans plusieurs domaines, notamment la typologie, les études comparatives, la linguistique historique et les représentations phylogénétiques.

Le projet tel qu'il a été conçu est évidemment loin d'être réalisé. Prévu pour une durée de deux ans, il a effectivement démarré il y a un peu plus de six mois. Cependant, beaucoup a été fait et des résultats ont déjà été obtenus :

- Les étapes initiales du projet (à l'exception du travail de terrain) ont pu être menées à bien : analyse des sources disponibles, inventaire des lacunes de la documentation.
- La grande majorité des données lexicales disponibles sur les langues BAK a été numérisée (saisie, harmonisation graphique).
- Ces données ont été organisées en une base de données comparative.
- Divers outils d'aide à la comparaison ont été mis au point, notamment :
 - un module d'analyse phonologique
 - un module de statistiques phonologiques
 - un module d'organisation des séries comparatives
 - un module d'organisation des séries de correspondances phonétiques.

Grâce à ces outils, il a été possible de dégager plusieurs centaines de séries comparatives lexicales pour les langues de l'ensemble Joola. Ces séries comparatives ont à leur tour permis d'établir l'inventaire des correspondances phonétiques régulières. Ces dernières sont en cours d'analyse pour permettre l'élaboration du scénario évolutif de ces langues, et l'établissement du proto-lexique.

En outre, le développement de ces outils a permis de poser les bases techniques d'un projet plus ambitieux : le *Corpus de Référence des Lexiques des Langues d'Afrique* (RefLex). Pour ce projet, que j'ai l'intention de soumettre à l'ANR l'année prochaine, j'ai commencé à prendre des contacts avec diverses équipes de recherche en Europe : le Musée Royal de l'Afrique Centrale à Tervuren, Belgique ; les universités de Bayreuth, Cologne, Hamburg, Mainz, en Allemagne ; l'université de Naples - l'Orientale, en Italie ; le laboratoire CNRS Dynamique du Langage, à Lyon. Il

s'agit de proposer à la communauté scientifique un corpus lexical pour les langues d'Afrique, qui réponde aux exigences suivantes :

- inclure les données publiées ou de première main
- proposer une transcription homogène et unifiée
- proposer un accès aux documents originaux
- fournir des outils de traitement et d'analyse des données

Les retombées scientifiques attendues sont importantes, et présentent deux aspects : d'une part, une évolution des pratiques, avec la possibilité pour l'ensemble de la communauté de procéder à des vérifications sur les données originales ; d'autre part, un véritable changement d'ordre de grandeur pour des champs disciplinaires tels que la linguistique historique et comparative, l'étude des contacts de langues, la typologie sémantique, la phonologie quantitative, etc.

Au-delà (mais dans la continuité) du projet de départ sur les langues BAK, et parallèlement au projet RefLex, mon intention est d'utiliser les résultats acquis et les méthodes mises en place pour enfin éclaircir la question de l'unité des langues de la famille Atlantique du phylum Niger-Congo. Ce programme, que j'ai présenté au concours externe pour le grade de CR1 et qui m'a valu d'être classé deuxième, est détaillé dans le document joint.

Annexe

Organisation de manifestations internationales Communications à des colloques

Organisation de manifestations :

2007, 25 septembre, Paris : Atelier *Typologie des langues africaines*, événement adjacent au 7^e colloque de l'*Association for Linguistic Typology*. Conférenciers invités : Zygmunt FRAJZYNGIER, Laura DOWNING, Tom GÜLDEMANN, Konstantin POZDNIAKOV.

2004, 11-14 octobre, Villejuif : *Proto-Niger-Congo Conference*. Atelier organisé avec Konstantin POZDNIAKOV et Sergei STAROSTIN, dans le cadre du projet *Evolution of Human Languages* du *Santa Fe Institute*.

Communications :

2008 : **invité** au Workshop *Documenting convergence and diversity – Mande and Atlantic languages in contact*, Londres, SOAS, 6-9 septembre.

2008 : **invité** à l'*Ecole d'été du Cercle Linguistique de l'INALCO*, Porquerolles, 25 au 30 mai.

2008 : **invité** au Workshop *Nouvelles perspectives en linguistique historique*, Lyon, DDL, 12-15 mai.

2008 : Lyon – Laboratoire DDL, 1^{ère} Journée d'études sur les langues atlantiques, 14 juin. *Titre à déterminer*.

2007 : Paris, ALT 7 (7^e colloque de l'*Association for Linguistic Typology*), 25-28 septembre.

– (avec Martine Vanhove, Karine Duvignau et Bruno Gaume) *From lexical semantics to semantic networks: a computational and typological study*

- 2007 : Leiden (Pays-Bas), 37th Colloquium on African Languages and Linguistics, 27-29 août 2007
 – *Une consonne orpheline dans les langues atlantiques.*
- 2007 : Leipzig (Allemagne), 28 juin. Représentant du LLACAN à l'atelier *Towards a Comprehensive Language Catalogue* organisé par Martin Haspelmath au Max Plank Institute for Evolutionary Anthropology.
- 2007 : Hambourg (Allemagne). 'Keynote Speaker' **invité** au colloque 'The Atlantic branch of Niger-Congo: Genetic or typological unit?', 17-18 février 2007
 – *Noun classification and agreement systems in Atlantic*
- 2006 : Londres (Royaume-Uni), International Conference on Bantu Grammar : Description and Theory, 22-22 avril 2006
 – *The Bantu-like patterns of the Bijogo verbs.*
- 2006 : Porquerolles (France), Présentation des travaux du PICS 'L'attribution de la qualité dans les langues africaines' (LLACAN, Univ. de Bayreuth), 4-8 octobre 2006
 – *Closed adjective classes and primary adjectives in African Languages*
- 2006 : Lyon – Laboratoire DDL, 1^{ère} Journée d'études sur les langues atlantiques, 8 février.
 – *Adjectifs et qualification dans les langues atlantiques*
- 2005 : Leiden (Pays-Bas), 35th Colloquium on African Languages and Linguistics, 29-31 août 2005
 – *10000 pronoms africains : la base de données du LLACAN.*
- 2004 : Villejuif – LLACAN, Proto-Niger-Congo Conference, 11-14 octobre 2004
 – *Présentation générale des langues atlantiques*
 – *La statistique et les perspectives de reconstruction du système phonologique du Niger-Congo*
 – *Perspectives d'élaboration des bases de données Niger-Congo*
- 2003 : Tervuren (Belgique), Musée Royal de l'Afrique Centrale, Séminaire de Langues et Cultures africaines, 7 novembre 2003 : *Le bijogo : la plus bantoue des langues atlantiques ?*
- 2003 : Lisbonne (Portugal), Annual meeting of the Language Typology Resource Center, 25-27 septembre 2003
 – *Présentation de la base de données "Pronoms" du LLACAN*
- 2003 : Leiden (Pays-Bas), 33rd Colloquium on African Languages and Linguistics, 25-27 août 2003 :
 – *Sémantique multidimensionnelle des classes nominales : l'exemple du bijogo.*
- 2001 : Berkeley (Etats-Unis) 32nd Annual Conference on African Linguistics, 22-25 mars 2001 :
 – *Où classer le bijogo?*
 – *Les systèmes de marques personnelles dans les langues atlantiques (avec K. POZDNIAKOV).*