

# LABORATOIRE LIGERIEN DE LINGUISTIQUE

Université d'Orléans / Université de Tours

LLL EA 3850 Institut de Linguistique Française (CNRS)



## Plan du dossier

RESUME / SUMMARY

1. Le laboratoire : une fusion exemplaire

ORGANIGRAMME

2. Décrire les langues à partir de corpus oraux

3. Impliquer la recherche dans la formation

4. Du local à l'international

5. Une gestion au service des missions

6. Production du laboratoire et activités des équipes

QUELQUES ELEMENTS DU BILAN SCIENTIFIQUE

15 PUBLICATIONS REPRESENTATIVES

PUBLICATIONS (ESLO, LANGUES, PHONOLOGIE, SEMANTIQUE, DIDACTIQUE)

EQUIPES, AXES, ACTIONS

## RÉSUMÉ

Le Laboratoire Ligérien de Linguistique résulte de la fusion de deux laboratoires de linguistique des universités d'Orléans et de Tours au 1<sup>er</sup> janvier 2008. Les travaux menés concernent les langues d'enseignement (français, anglais, allemand, espagnol) et la description de langues africaines (ikwéré, sèmè, dagara, créoles portugais) et américaines (palikur, wayana, wayampi, chiac). L'accent mis sur la dimension orale des langues dans la constitution d'une linguistique de corpus et dans l'analyse des données (morpho-phonologie, sémantique et prosodie) a permis de dégager progressivement, dans le travail réalisé par les différentes équipes, un intérêt commun pour des méthodes et un domaine. Ces études trouvent leur prolongement dans les applications didactiques et dans la mise à l'épreuve des outils et instruments du traitement automatique des langues dans l'exploitation des données orales. Au nombre des résultats, on peut mentionner la constitution de la plus grande base sociolinguistique de français oral en ligne, les études sur le catalogage et le codage, l'articulation morphologie-phonologie de l'anglais et la diachronie, la spécialisation dans les études sur les créoles à base portugaise, la recherche cognitive, la didactique de l'écrit ou la publication de l'*Encyclopédie des Peuples et Cultures de Guyane*. Ces recherches s'appuient sur une politique de formation, de la licence au doctorat et sur des collaborations en Europe, en Afrique et en Amérique.

## SUMMARY

The LLL—Loire Linguistics laboratory was set up by merging Orléans and Tours's linguistics research units on January, 1, 2008. Research work has focused on taught languages (French, English, German and Spanish) and the description of African languages (ikwéré, sèmè, dagara, portuguese creole) et American ones (palikur, wayana, wayampi, chiac). Stress was put on the oral dimension of languages within the framework of the setting up of a corpus linguistics and the analysis of data (morpho-phonological, semantic and prosodic) thus enabling the different teams to reach a common interest in terms of method and research area in their work. These studies have led to didactic applications and to hands-on experience of CAT(computer assisted tools) tools in the treatment of oral data. Amongst results, it is worth mentioning the greatest online oral French sociolinguistic database, the studies on encoding and cataloguing, morphology-phonology for English and diachrony, specialization in Portuguese-based creole studies, cognitive research, Written didactics or the editing of the *Encyclopédie des Peuples et Cultures de Guyane*. The research is backed by a training-oriented policy from the third graduating year to the doctorate, and by European, African and American partnerships.

## 1. Le laboratoire : une fusion exemplaire

Le LLL est devenu, depuis le précédent contrat quadriennal, le laboratoire linguistique de référence en Région Centre, présent sur les deux universités. Il résulte du rapprochement, lors du précédent quadriennal, du **Centre Orléanais de Recherche en Anthropologie et Linguistique** (CORAL/EA 3850, Université d'Orléans) et de l'équipe **Langage et Représentations** (Université de Tours), équipe soutenue jusqu'alors par le Conseil scientifique de l'Université François Rabelais. La démarche était volontariste, le consensus sûr, les relations interpersonnelles bonnes et les projets solides. La qualité de la collaboration a fait de cette expérience d'un laboratoire régional, le seul jusqu'à présent au niveau du PRES, et de l'avis général, une réussite,

Au 1er septembre 2010, le LLL réunissait **30 chercheurs et enseignants-chercheurs** :

6 professeurs (+ 1 professeur émérite) de 7<sup>e</sup> et 11<sup>e</sup> sections, 4 à Orléans et 2 à Tours,

1 DR CNRS détaché en Guyane pour diriger la mission "Oyapock",

22 MCF, 16 à Orléans (dont 3 à l'IUFM) et 6 à Tours,

1 CR CNRS en rattachement temporaire (demande en cours d'instruction)

et aussi :

**27 doctorants** (dont 4 ATER et 3 allocataires),

5 MCF avec le statut d'associés, et une trentaine de collègues, PRAG, maîtres du premier et du second degré, professionnels, chercheurs en retraite...

Composition du bureau

Directeur : Gabriel Bergounioux, PR de linguistique, Université d'Orléans

Directeur-adjoint : Jean-Michel Fournier, PR d'anglais, Université de Tours

François Nemo (directeur-adjoint pour le site d'Orléans, Master), Antonia Cristinoi (Budget), Emmanuel Schang (programmes) à Orléans, Sylvester Osu (Master), Fabienne Toupin (Budget) et Nathalie Garric (programmes) à Tours.

Au cours des années passées, le LLL a connu le départ de 8 EC :

- 4 PR (1 en 14<sup>e</sup> section à Orléans, retraite, 1 en 11<sup>e</sup> section à Tours, retraite, et 2 en 12<sup>e</sup> section à Tours, mutation) ;

- 3 MCF (2 en 7<sup>e</sup> section, 1 à Tours, changement de laboratoire et 1 à Orléans, retraite, et 1 en 14<sup>e</sup> section à Orléans, retraite).

Il a accueilli dans le même temps 10 EC ou CR :

- 1 PR en 11<sup>e</sup> section à Tours,

- 8 MCF (2 en 7<sup>e</sup> section à Orléans, 5 en 11<sup>e</sup> section, 3 à Orléans et 2 à Tours, et 1 en 18<sup>e</sup> section),

- 1 CR CNRS (poste en cours de négociation avec la Direction scientifique et la section 34).

Un onzième EC (70<sup>e</sup>/7<sup>e</sup>) devrait être recruté au fil de l'eau d'ici décembre 2010.

Si l'évolution est positive et si les recrutements effectués correspondent dans l'ensemble à des programmes bien identifiés, les aléas des recrutements ont conduit à une rétraction de l'espagnol (à Orléans) et de l'allemand (à Tours) qui n'a pas encore été compensée par une politique de recrutement en sciences du langage correspondant aux besoins des départements d'enseignement.

La participation aux comités de sélection a permis de définir une politique de recrutement qui, en accord avec la politique conduite par l'université, a mis l'accent sur la contribution à la recherche.

## ORGANIGRAMME DU LLL

### UNIVERSITÉ DE TOURS

(ED SHS Tours)

Direction J.-M. Fournier (dir.adj.)

#### Bureau

Master S. Osu  
Programme N. Garric  
Budget F. Toupin  
Doctorant M. Martin

#### Programmes

ESLO/Variling  
Analyse discours N. Garric  
Lg d'Afrique S. Osu  
Lg de Guyane  
Morphonologie J.-M. Fournier  
Diachronie anglais F. Toupin  
GORDF  
PROHEMIO

#### Actions transversales

Ling. cognitive G. Col  
Rep. du sens S. Osu  
Phonologie J.-M. Fournier

### UNIVERSITÉ D'ORLÉANS

(ED SHS Orléans)

G. Bergounioux (dir.)

F. Nemo (dir.adj. Orléans)  
E. Schang  
A. Cristinoi  
L. Hriba

O. Baude

J.-L. Rougé  
F. Grenand

J. Lafont  
S. Fournié

F. Nemo

G. Bergounioux

Le bureau, réélu chaque année, assure la gestion du laboratoire entre deux assemblées générales dont la périodicité est annuelle. L'assemblée générale, dont sont membres tous les EC et chercheurs en rattachement principal, approuve les comptes et définit la politique du laboratoire. Les doctorants et les personnes associées aux programmes de recherche, en France et à l'étranger, sont invités à l'AG. Les doctorants élisent leur représentant au bureau.

## 2. Décrire les langues à partir de corpus oraux

Le LLL a fondé ses recherches sur la description des langues, qu'elles correspondent à des cursus d'enseignement au sein des établissements de rattachement (français, anglais, allemand et espagnol) ou à des engagements sur différents terrains en Afrique (kriol, ikwere, dagara, sèmè...), en Amérique du sud (palikur, wayana, wayampi) et du nord (chiac).

Le projet scientifique, au-delà des différences de structure entre langues et de la disparité de la couverture spatiale, s'est progressivement centré sur les méthodes d'investigation. La caractéristique première des langues, avant toute consignation graphique, tient à ce que l'oralité s'avère inséparable d'une instabilité des productions sonores et d'une perte obligée d'information au moment de la conversion en écriture. C'est pourquoi, sur des langues diverses et avec des approches variées, la constitution des données a retenu l'attention des équipes, des modalités de l'enquête sur le terrain à l'exploitation de sources écrites, en passant par les questions de nature juridique ou archivistique. Réciproquement, les effets de l'écriture sur la notation des langues, dans la traduction prosodique des entrées lexicales ou dans les modalités d'apprentissage rédactionnel, ont alimenté les recherches.

Si le programme ESLO a été prototypique dans ce travail, l'ensemble des équipes et de nombreux projets ont concouru à forger une expertise dans l'un des domaines les plus actifs de la linguistique à l'échelle internationale, la linguistique de corpus, un secteur où la France contribue activement à la dynamique du champ.

Les ressources disponibles, centralisées par le CRDO (Centre de Ressources pour la Description de l'Oral / Aix-en-Provence) et par le TGE-ADONIS, ont été accrues par la participation du laboratoire à ce projet collectif. Pour assurer un traitement qui prenne en compte les attentes et les requêtes, de nouveaux outils et de nouveaux instruments, de nouvelles ressources ont été constituées.

Sont énumérées ci-dessous quelques initiatives et différents résultats dans ce domaine, en relation avec les phases d'études :

***Constitution de corpus écrits*** : archives de l'enquête (littérature grise, documents de travail), base de données sur les témoignages écrits de langues parlées (chiac), recueil de productions scolaires en génétique du texte...

***Constitution de corpus oraux*** : programme ANR « Variling », programmes PANGLOSS (langues en contact) et RADICEL-K (anthropologie linguistique du KénéDougou), collectes d'enregistrement sur le terrain en Afrique (Guinée-Bissau, Sénégal, Burkina, Nigéria...) et en Amérique (Guyane)...

***Opérations de catalogage et de codage*** : collaboration AnthroNet avec les Musées et centres de recherche en anthropologie, colloque international CAT-COD sur la TEI et l'Open Language Archives Community (OLAC)...

***Protection juridique*** : direction du *Guide des bonnes pratiques dans l'utilisation des corpus oraux*, élaboration des documents sur l'exploitation et l'utilisation des données à destination des témoins, conventions de dépôt.

**Transcription** : élaboration d'un système de transcription du palikur, définition des alignements sur signal (en requête sur la Toile comprise), analyse d'erreurs en multi-transcription.

**Comparabilité des corpus** : formatage des données pour la comparaison des enregistrements et des locuteurs, principes d'interopérabilité, études en contrastif des entités dénommantes.

**Traitement informatique** : étiquetage morpho-syntaxique par apprentissage automatique avec des CRF (en collaboration avec le Laboratoire d'informatique d'Orléans), reconnaissance automatique des entités nommées et dénommantes par cascade de transducteurs, extraction automatique du lexique (palikur), détection automatique des métaphores, repérage automatique de phonologie prosodique des connecteurs, analyse automatique morpho-phonologique de l'anglais.

**Analyse linguistique** : études phonologiques, dénomination, sémantique cognitive, définition des proximités discursives, coréférence, analyse de discours, parémiologie ibérique...

**Interface anthropologique** : épistémologie des identités (colloque de Tours, 2008), programme d'études sur El Rebollar (Espagne), l'Oyapock (Guyane), le Kéné Dougou (Burkina)...

**Conversion grapho-phonologique** : encodage des bases lexicales de l'anglais, construction d'une analyse des productions écrites...

**Applications didactiques** : collaboration en cours en lien avec l'IUFM Val-de-Loire, le CIEP et la MSHS de Poitiers sur l'écrit scolaire et universitaire.

L'expertise du LLL, dans son domaine de compétence, a trouvé son prolongement dans les deux domaines d'application de la linguistique, en TAL et en didactique. Un transfert de compétence dans la définition de méthodes robustes et pérennes de fabrication et de sauvegarde de témoignages de langues parlées dans des conditions difficiles a été assuré lors de missions au Burkina et par la formation de chercheurs au Burkina, au Cap-Vert et au Nigéria (en collaboration avec l'IFAN). La production de nouveaux savoirs sur les langues a été réalisée soit en constituant une archive de langues peu ou pas décrites qui initie une description et préserve les possibilités d'une étude approfondie ultérieurement, soit en apportant de nouveaux matériaux sur des langues déjà connues dans des formats qui en facilitent la transmission et l'exploitation. On peut citer, dans le premier cas, les recherches menées sur les créoles à base portugaise (Cap-Vert, Guinée Bissau, Sao Tome et Principe) pour lesquels les matériaux recueillis doivent aboutir à une synthèse à paraître dans la collection "Les Langues du Monde" de la Société de Linguistique de Paris ; dans le second cas, le recueil des témoignages réalisé en Espagne par le groupe PROHEMIO.

### 3. Impliquer la recherche dans la formation et la formation dans la recherche

Le LLL est présent sur les deux universités du PRES Centre Val-de-Loire : Orléans et Tours. C'est le premier laboratoire de la région à avoir étendu ses activités sur les deux établissements, dans cinq composantes (UFR Lettres et Langues à Tours, UFR Lettres Langues et Sciences Humaines, IUT et IUFM à Orléans, plus la mission CNRS de Guyane). Ce regroupement a transcendé la division en départements puisque se trouvent réunis des collègues de sciences du langage, d'anglais, d'allemand, d'espagnol et de techniques de communication, partageant le même objet et les mêmes méthodes. Grâce à ces intérêts communs, présent sur deux sites, trois institutions et six sections du CNU, le LLL n'a pas connu de véritable difficulté pour trouver sa cohérence.

Au niveau de la **licence**, les EC du LLL animent :

- la licence SDL d'Orléans et, partiellement, de Tours, comprenant des orientations différenciées, à partir de la L3 notamment, en « FLE », « FLM », « TAL » et « Communication »,
- la composante linguistique des licences d'anglais (Orléans et Tours) et d'allemand (Tours).

En moyenne, au cours du contrat, ce sera environ quatre-vingts étudiants par an qui auront été titulaires d'une licence SDL, une formation mise en place à Tours à l'occasion du précédent quadriennal. Il est à noter que dans la partie sociolinguistique des enseignements et dans les parties didactiques, il y a eu non seulement une présentation des travaux effectués dans le LLL mais association d'un nombre important d'étudiants à des enquêtes et à des études dans l'école.

Au niveau du **Master**, dans la même démarche qui avait abouti au rapprochement des laboratoires, il avait été convenu qu'une même maquette serait présentée dans les deux universités en SDL. Ce Master, dont les effectifs cumulés ont pu dépasser pour les deux années et les deux sites 150 étudiants, s'est décliné à Orléans avec :

- une spécialité en « linguistique avancée et applications aux STIC », correspondant à l'ancienne spécialité recherche et à une orientation dans les domaines du TAL (ingénierie linguistique et traitement de la communication),
- une spécialité « linguistique et didactique », comprenant une partie FLE et une partie FLM, celle-ci en collaboration avec l'IUFM, et une option pour la conception de didacticiels unique en France.

A Tours, seule la partie « linguistique avancée et applications » a été développée.

Dans les deux établissements, les linguistes anglicistes participent aux Masters de leur département, en particulier dans les métiers de l'enseignement et de la formation. A Orléans, deux collègues du LLL animent le parcours traduction du Master Lettres Langues et Affaires internationales. Il existe des collaborations avec le Master « Compétences complémentaire en informatique » assuré par l'UFR Sciences et avec le Master « Conseil politique et communication » assuré par l'UFR Droit Economie Gestion.

Le LLL est par ailleurs devenu le laboratoire de référence de l'IUFM pour ce qui concerne la formation des professeurs des écoles qui a été confiée à cette composante. Cette filière s'articule avec une présence continue dans le domaine de la formation continue (plan de formation continue des personnels) et avec des travaux conduits en collaboration avec les chercheurs de l'établissement, en sciences de l'éducation et en psychologie.

Au niveau **doctoral**, le LLL s'est trouvé confronté à la question de la diversité des sujets de thèse et à l'étréitesse des personnels HDR à même d'encadrer les doctorants. Si, de facto, le LLL s'est trouvé très impliqué dans l'animation doctorale dans les universités puisque le directeur de l'ED SHS d'Orléans était en même temps le directeur du laboratoire, que le LLL a bénéficié, sur les deux sites, d'une allocation par an en moyenne, qu'un doctorant a figuré au bureau de l'Association des Docteurs et Doctorants en SHS, les nécessités d'une ouverture maximale des compétences a conduit à combiner quatre types de préparation à la rédaction et à la soutenance de la thèse.

D'abord, les étudiants ont profité au mieux des formations communes de l'ED, qu'il s'agisse de cours de langue ou de présentation d'outils informatiques, en particulier de LATEX devenu aujourd'hui la référence en matière de PAO, et de statistiques.

Puis, en linguistique, les étudiants ont été associés aux séminaires du laboratoire et des équipes afin de se familiariser, en même temps qu'avec les contenus de la recherche, avec un format de présentation. Ces exposés trouvaient leur prolongement dans une participation aux activités de recherche conduites à l'intérieur des axes ou des équipes.

Ensuite, afin de faire profiter aux mieux les doctorants du travail en cours, il a été convenu que les heures de formation prévues par le contrat doctoral seraient effectuées au maximum dans des séminaires doctoraux spécialisés (en phonologie, en linguistique berbère...) accessibles dans les universités voisines et complétées par la participation à des journées d'études, conférences et colloques.

Enfin, chaque étudiant bénéficie d'un soutien individuel de la part de son directeur de thèse, de l'ensemble des enseignants-chercheurs du laboratoire et bien sûr des autres doctorants puisque des bureaux et des ordinateurs sont mis à leur disposition à l'université, de façon plus avantageuse à Orléans qu'à Tours où la question des infrastructures immobilières continue à se poser.



## 4. Du local à l'international

Au plan **local**, le LLL est associé à la politique de recherche de l'université à travers ses différentes instances. A Tours, où il figure au nombre des équipes constitutives de l'ED SHS, il a été prévu une intégration, qui s'étendra à l'ensemble de la structure, dans la Maison des Sciences de l'Homme reconfigurée qui se met en place à l'occasion du nouveau contrat. A Orléans, cette anticipation d'une structure fédérative, qui devrait avoir son répondant sur site, a conduit le LLL à figurer comme l'une des trois équipes du premier cercle du Centre de Ressources et d'Innovation sur les Territoires (CRITERR) avec le LEO (Economie) et le CEDETE (Géographie). Le LLL y a notamment inscrit une recherche sur les pratiques linguistiques en région Centre (constitution de corpus oral, étude sur la diffusion des langues d'immigration, cellule de réflexion sur la simplification des documents administratifs, projet de phonothèque régionale) et sur la coopération Nord-Sud (programme conjoint avec les chercheurs du Burkina et les géographes du CEDETE).

**Au-delà de la région**, une organisation a été esquissée à partir de la mise en place d'une fédération de laboratoires associant au LLL une équipe INSERM de Tours ("Langage et cerveau"), le CeReS de l'Université de Limoges et le FORELL-A de l'Université de Poitiers. Une première étape, au-delà de la construction d'une structure fédérative, qui n'a pas donné les résultats escomptés (voir rapport sur FELINCO), a été l'intégration du LLL dans le CERLICO, un groupement associatif où se retrouvent nombre de linguistes de Poitiers, La Rochelle, Limoges, Nantes, Angers, Le Mans, Rennes, Brest, Lorient-Vannes, Caen et Bordeaux. Centré sur l'organisation d'un colloque annuel, édité aux PUR, et une publication électronique, CORELA (voir site), dont le directeur est membre du LLL et dont la maintenance s'effectue à Poitiers, le CERLICO représente l'esquisse d'une collaboration large entre équipes à une échelle interrégionale.

A ce rapprochement géographique correspondent des insertions thématiques au plan **national** :

- en linguistique française (Institut de Linguistique Française, fédération de recherche du CNRS), en linguistique cognitive, dans la collaboration entre le travail conduit à CORELA et le LATTICE (voir document annexe)
- en créolistique (dans l'AICBLPE, voir site),
- dans le domaine des études anglaises (colloques ESP à Orléans en 2008, colloque Linguistique diachronique de l'anglais en 2009 à Amiens).

C'est une politique qui est appelée à se développer, en particulier par des conventions associant le LLL aux laboratoires aréaux dont les compétences sont un appui décisif. Des discussions ont été amorcées, pour le moment informelles, avec le LLACAN.

Dans les secteurs disciplinaires, le LLL est très présent dans les domaines de la sémantique (*RSP*, *CORELA*, rencontres de sémantique et pragmatique à Gabès 2009) et de la phonologie (organisation des colloques CERLICO "Transcrire, Ecrire, Formaliser" à Tours en 2009, à Orléans en 2010, des colloques du Réseau Français de Phonologie à Orléans en 2009 et à Tours en 2010). Il entend l'être plus encore dans la structuration du réseau qui sera en charge de la mise en place d'un corpus de référence du français parlé, en partenariat avec la BnF, les chercheurs en France et dans le monde qui seront appelés à y contribuer, dans les structures qui en préfigurent la création : TGE ADONIS, CRDO, DGLFLF et, bien sûr, l'ILF.

Au niveau **international**, le LLL entretient des rapports structurés, non exclusifs de relations avec des chercheurs d'autres pays, avec les universités ou instituts suivants, avec précision du domaine ou du programme concerné.

Dans l'enseignement :

En master :

Jagellone (Pologne/ERASMUS),  
Cluj-Napoca (Roumanie/ERASMUS),  
Gabès (Tunisie),  
Praia (Cap-Vert),  
Ibadan (Nigéria),  
Libreville (Gabon),  
Nankin (Chine).

En doctorat :

Beyrouth (Liban),  
Cluj-Napoca (Roumanie),  
Ouagadougou (Burkina),  
Praia (Cap-Vert),  
Le Caire (Egypte),  
Naplouse (Palestine/Israël).

En recherche :

Louvain-la-Neuve (Belgique) – ESLO + GORDF,  
Salamanque (Espagne) – PROHEMIO,  
Coimbra (Portugal) – Créoles & Lg d'Afrique,  
Sao-Paulo (Brésil) – Créoles & Lg d'Afrique,  
Campinas (Brésil) – Lg d'Amérique,  
Moncton (Canada) – Lg d'Amérique,  
Ouagadougou (Burkina) – Créoles et Lg d'Afrique.

## 5. Une gestion au service des missions

L'organisation du LLL est décentralisée par site et par équipes afin de simplifier le processus décisionnel. Chaque équipe dispose de ses propres lignes de financement, soit par contrat, soit par abondement sur la dotation récurrente, les dépenses exceptionnelles étant couvertes par les fonds communs. A noter que les doctorants bénéficient de crédits particuliers pour toutes les actions de formation. Toutes ces dépenses sont visées avant exécution par le directeur (Orléans) ou le directeur-adjoint (Tours) et par la responsable financière du laboratoire sur site. La supervision globale, incluant les équipes et les sites, intervient a posteriori lors de l'assemblée générale du laboratoire qui réunit l'ensemble des personnels et à qui est soumis le bilan financier.

### AG et bureau

L'AG est l'instance décisionnelle. Elle se réunit au moins une fois par an, durant une journée, afin de faire le point administratif (élection annuelle du bureau), financier et scientifique. Entre ces AG (les statuts du labo prévoient explicitement qu'elle peut être convoquée à la demande autant que de besoin, sans que cette disposition ait été nécessitée jusqu'à présent), le bureau se réunit quatre fois par an. Il comprend huit personnes, quatre pour chacun des sites :  
le directeur (Orléans) ou le directeur adjoint (Tours),  
le gestionnaire financier,  
le responsable des études doctorales et de la partie recherche du Master,  
un représentant des doctorants.

C'est au bureau qu'il incombe d'appliquer les décisions prises en AG, de coordonner les activités sur les deux sites, d'assurer les relations entre les équipes et de prendre les décisions dont il rendra compte devant l'ensemble des collègues. Si l'interaction et le caractère serein des relations ont rendu faciles l'application des décisions et les échanges, en revanche, la distance et les différences de mode de fonctionnement entre les deux établissements ont parfois compliqué la tâche de ceux qui se sont engagés au service du collectif, sans répercussion majeure sur l'activité quotidienne si l'on excepte les événements du premier semestre 2009.

### Immobilier et financements

Concernant les aspects matériels, le LLL bénéficie d'un triple hébergement : à Tours sur le site des Tanneurs, à Orléans dans l'UFR Lettres et à l'IRD. Comme EA inscrite au précédent contrat avec une évaluation A, le LLL a bénéficié de financements récurrents de la part de ses universités de rattachement, financements qui ont été en augmentation sensible ces dernières années (mais un accroissement en partie amoindri par la fin du BQR). Le fléchage de crédits pour FELINCO, en particulier pour des collaborations avec Poitiers, a représenté un effort significatif supplémentaire de l'établissement. Les seuls autres crédits récurrents du laboratoire proviennent de CRITERR mais l'absence de prévision sur leur disponibilité et le caractère tardif de leur allocation aux équipes, joint à la nécessité d'une dépense précipitée, ont rendu cette ressource plus difficile à utiliser, en particulier pour que l'affectation soit intégrée au déroulement normal des opérations financières internes. A Tours, une autre politique d'allocation des ressources a abouti aussi à un accroissement des fonds disponibles pour la recherche mais leur insuffisance pour des projets de plus grande ampleur a conduit à un transfert d'Orléans vers Tours.

Une deuxième ligne de financement est liée à des programmes et des projets :

- nationaux : ANR Variling, PEPS, contrats DGLFLF et soutien de l'ILF
- interuniversitaires : CO2, programme Tours-Poitiers sur la parole

- régionaux : PANGLOSS, RADICEL-K

Une troisième ressource accompagne des actions ponctuelles : appui à des colloques apportés par l'établissement et les collectivités territoriales ou action à caractère exceptionnel. Au nombre de ces opérations figure un crédit de 30.000 Euros, complétant des ressources propres du laboratoire, pour le traitement par une entreprise (société Ares) de la mise en ligne de corpus oraux, une opération technique complexe que les services informatiques de l'Université ne pouvaient assumer.

### **Une difficulté : l'absence de personnel administratif et technique**

On est, avec ce cas de figure, en présence de la plus grande difficulté à laquelle se soit trouvé confronté le LLL : l'absence de personnel pour les tâches de gestion et de secrétariat. Dans les derniers mois, un CDD précaire a été mis à la disposition des quatre laboratoires de l'UFR LLSH, un soutien tardif et temporaire qui n'est pas proportionné aux besoins et qui requiert un nouveau calcul des moyens à l'échelle de l'établissement. Enfin, le Centre de Ressources Informatiques de l'établissement n'a pu contribuer à l'activité du laboratoire, en particulier pour la mise en place d'un site pour le laboratoire à l'intérieur du domaine de l'université, imposant un appel d'offres auprès de prestataires extérieurs pour pallier ce problème.

Ce sont aujourd'hui les ressources humaines qui représentent le principal obstacle au déploiement des recherches et des résultats obtenus, comme le montre la part très importante, pour un labo de SHS, consacrée aux prestations et aux CDD dans le compte financier.

## 6. Productions du laboratoire et activité des équipes

Le LLL a participé activement à la production scientifique globale en matières de publications. Le détail est fourni pour chacun des sites, à Orléans par équipe et 0 trous par axe et action transversale. Poursuivant sur un mode d'organisation établi dans l'équipe «Langages et Représentation», Tours a souhaité conserver la distinction axe/thématique transversale. Une restructuration en équipes, avec poursuite d'une action transversale sur site, a reconfiguré le mode de fonctionnement du laboratoire pour le prochain contrat, unifiant les pratiques dans les deux établissements. L'ouverture du programme de recherche à la BnF marquera une extension des formes de travail collaboratives (séminaires, journées d'études...).

### *ORLEANS* (par équipes)

Corpus variationniste du français / ESLO  
Langues et cultures de Guyane  
Langues d'Afrique et créoles  
Recherches en didactique du français  
PROHEMIO (culture orale de l'Espagne)

### *TOURS* (par axes)

Morpho-phonologie de l'anglais  
Linguistique cognitive  
Analyse de discours  
Langage et traitement de l'information  
Diachronie de l'anglais  
Action thématique : Dénomination + Temporalité

### *ORLEANS & TOURS* (actions transversale)

Corpus oraux / Variling  
CAT COD  
CORELA  
RSP  
Réseau Français de Phonologie

Ces productions ont pris la forme particulière des ressources requises par les études linguistiques aujourd'hui, c'est-à-dire de corpus, en tant que collecte de données orales, identifiées, situées et hiérarchisées, qui ont fait l'objet d'un travail de catalogage, de transcription et de relecture et qui sont associées à des métadonnées, accessibles de façon différenciée, pour la recherche scientifique. Le temps nécessaire pour leur constitution, qui est inséparable de la construction des objets scientifiques en linguistique, se matérialise dans la disponibilité réelle du nombre d'heures exploitables répondant à un cahier des charges strict, celui présenté dans l'appel d'offres Variling en treize points. Si l'on convient que la définition proposée pour qu'un corpus soit validé en tant qu'outil (et non seulement comme une ressource) correspond au niveau d'exigence déclaré par ESLO, le LLL peut se targuer de disposer aujourd'hui de la plus grande ressource exploitable en ligne en matière de français oral.

D'autres corpus sont en cours d'élaboration concernant les langues de migration (programme Langues en Contact à Orléans), le sèmè (Burkina), le palikur (Guyane), le kriol (Cap-Vert)... Des

bases de données sur l'écrit ont été réalisées dans le cadre du programme didactique et dans la correspondance graphie/phonie à partir des ressources électroniques de plusieurs dictionnaires de l'anglais.

### **Diffusion de la recherche**

L'activité de la recherche comprend aussi la circulation de la recherche à travers la publication de supports écrits sur papier et/ou CD :

*Revue de Sémantique et Pragmatique, Cahiers du PROHEMIO, Encyclopédie des Langues et Cultures de Guyane,*

et de ressources en ligne :

*CORELA, Revue de l'AICBLPE, ESLO, colloques et journées d'études à Tours,*

de colloques :

7<sup>e</sup> colloque du PROHEMIO (El Rebollar, 2007),  
Construction d'identité et processus d'identification (Tours, 2007),  
CATalogage-CODage (Orléans, 2008),  
Anglais de spécialité / English for Specific Purpose (Orléans, 2008),  
Diachronie de l'anglais (Amiens, 2008),  
8<sup>e</sup> PROHEMIO (El Rebollar 2008 et 2009),  
5<sup>e</sup> Rencontres de Sémantique et Pragmatique (Gabès, 2009),  
Déconstruction et reconstruction : la pensée de J.-M. Zemb (Tours, 2009),  
CERLICO (Tours 2010 - Orléans 2011),  
Réseau Français de Phonologie (Orléans, 2010 et Tours, 2011)...

et de journées d'études :

Analyse linguistique des corpus discursifs (Tours, 2008)  
Corpus oraux et muséographie (Orléans 2009),  
Terrains d'Afrique (Orléans 2009),  
Le style et sa formalisation (Tours, 2009)  
Grammaires créoles (Orléans 2010)...

Enfin, l'activité de recherche concerne avant tout des personnes, qu'il s'agisse des étudiants de Master ou des doctorants, même si, pour des raisons qui supposent un examen au cas par cas des dossiers, les délais de soutenance sont encore trop importants. Les thèses soutenues sur les corpus (prosodie et sémantique par G. Cloiseau et M. Petit, dynamique du changement linguistique par A. Chesneau), sur la phonologie (morphologie de l'anglais par V. Abascq), sur la sémantique (iconicité et propagande par M. Feugain, parémiologie arabe et berbère par A. Talmenssour, argumentation polémique par D. Vlad) indiquent les axes de développement de la recherche dans le laboratoire.

Au cours de ce contrat, un certain nombre de collègues ont enrichi leur dossier en sorte qu'une politique de soutenance de HDR doit être élaborée sur un plan collectif pour le prochain quadriennal. C'est l'un des aspects sur lesquels les résultats ne sont pas à la hauteur des espérances lorsque l'on mesure la production des chercheurs et leur âge à l'absence d'HDR au cours des trois années écoulées.

## Quelques éléments du bilan scientifique

- la description de langues peu ou pas décrites (dagara, ikwéré, sèmè, krio, chiac, palikur, wayana, wayampi...), une tâche à poursuivre dans les années à venir, l'étude d'états de langue non stabilisés, qu'il s'agisse de créoles, de langues mixtes (chiac) ou de langues en contact en situation de migration (LCO), ou bien de données non stabilisées des langues (disfluences, multiplicité des transcriptions)
- l'élaboration de corpus qui doivent être étendus et aboutir à de nouvelles exploitations, à des applications et à une généralisation des procédures définies à grande échelle (interopérabilité)
- la définition de règles juridiques et techniques d'exploitation (CAT-COD, OLAC...)
- l'interface prosodie/sémantique, que ce soit à l'échelle lexicale (avec la maquette d'un dictionnaire multimodal) ou discursive
- l'interface graphie/morpho-phonologie, en anglais et en français
- la sémantique des langues naturelles et la linguistique cognitive
- le traitement automatique, en particulier dans le développement des interfaces qui assurent une accessibilité aux données alignant signal et transcriptions pour une utilisation en analyse de discours, avec un développement spécifique en annotation automatique (entités nommées, dénommantes, parties du discours).
- les applications didactiques (apprentissage de l'écrit, conception de didacticiels). Le détail des résultats se trouve exposé dans les parties correspondant aux programmes et aux chercheurs

## Bibliographie sélective de 15 titres représentatifs

01. Baude, O. coord. (2006) *Corpus oraux, guide des bonnes pratiques*, Paris et Orléans, Editions du CNRS et Presses Universitaires d'Orléans.
02. Bergounioux, G., *Bergounioux M., Nguyen N. & Wanquier S.* dir. (2007) « Mathématiques et Phonologie », *Mathématiques et Sciences Humaines* 180.
03. Bourdier, Ph. (2008) Un grand écran pour les Lettres : le cinéma et l'enseignement du français, « Nouvelles pédagogies », éd. L'Harmattan, Paris, 313 p.
04. Cadiot, P. & Visetti, Y.-M. (2006) *Motifs et proverbes, Essai de sémantique proverbiale*, Paris, Presses Universitaires de France.
05. Col, G. (2007) « Windowing the future. The cognitive operation of 'windowing' in the study of future time evocation », *From Gram to Mind: Grammar as Cognition*, J.-R. Lapaire, J.-R. et al. (eds), Bordeaux, Presses Universitaires : 323-341.
06. Fournier, J.-M. (2010) *Manuel d'anglais oral*, Paris : Ophrys
07. Garric N. & Légise I. (2007) « Aspects syntaxiques et discursifs d'un français parlé des médias : le discours d'information télévisé », in *Actes du Colloque international sur le Français Parlé des Médias* (8-12 juin 2005, Université de Stockholm) : 243-258.
08. Grenand, F. & Camargo, E., Capiberibe, A., Cristinoi, A., Gallois, D., Grenand, P. « Encyclopédies palikur, wayana, wayâpi : langue, milieu et histoire », fascicule 0 de l'*Encyclopédie des peuples de Guyane*, Paris/Orléans, CTHS-P.U.O., 165 p.
09. Iglesias Ovejero, A. et al. dir. (2008) *Cahiers du P.R.O.H.E.M.I.O.* 9 et 10, *Memoria histórica – Historia oral : guerra, postguerra y emigración en El Rebollar y otras partes*. Actas del VII Colloque PROHEMIO (20-22 juillet 2007), P.U.O., 406 pp. et 458 pp.
10. Lafont-Terranova, J. & Colin, D. (dir.) (2006) *Didactique de l'écrit : La construction des savoirs et le sujet-écrivain*. *Actes de la journée d'étude IUFM d'Orléans-Tours (mai 2005)*, Namur, CEDOCEF /P.U.N.
11. Nemo, F. (2006) « Discourse words as morphemes and as constructions ». In *Approaches to Discourse Particles* (Studies in Pragmatics, Volume 1), Fischer, K. ed., Elsevier Science : 415-448.
12. Osu, S., Col, G., Garric, N. & Toupin, F. (2010) *Construction d'identité et processus d'identification*. Berne : Peter Lang.
13. Rougé, J.-L. & Schang, E. (2006) « The origin of the liquid consonant in Saotomense Creole », *The Structure of Creole words*, Niemeyer-Verlag, Tübingen : XX-XX.
14. Serpollet, N. (2008) « A Study of the Mandative Subjunctive in French and its Translations in English: A Corpus-Based Contrastive Analysis », *Incorporating Corpora – The Linguist and the Translator*, Anderman, G. & Rogers, M. (eds), Clevedon, Multilingual Matters : 187-216.
15. Toupin, F. (2008) « Des phénomènes de pidginisation et de créolisation en moyen-anglais » *Un espace colonial et ses avatars. Naissance d'identités nationales : Angleterre, France, Irlande (V<sup>e</sup>-XV<sup>e</sup> siècles)*. Paris, PUPS : 179-201



## Productions majeures par domaine

### ESLO

**Abouda, L. & Baude, O.** (2007) « Constituer et exploiter un grand corpus oral : choix et enjeux théoriques. Le cas des Eslo », *Corpus en Lettres et Sciences sociales. Des documents numériques à l'interprétation*, Rastier, F. & Ballabriga M. dir, Actes du XXVII<sup>e</sup> Colloque d'Albi, Toulouse, Presses Universitaires du Mirail : 161-168.

**Baude, O.** coord. (2006) *Corpus oraux, guide des bonnes pratiques*, Paris et Orléans, Editions du CNRS et Presses Universitaires d'Orléans.

**Baude, O.** (2008) « Le droit de la parole », *Données orales, les enjeux de la transcription*, Bilger, M. (ed), Presses universitaires de Perpignan, p 23-34.

**Bergounioux, G.** (2010) « Mai 68 vu d'Orléans », Actes du 2<sup>e</sup> *Congrès Mondial de Linguistique Française* organisé par l'ILF-CNRS, Nouvelle-Orléans, 12-16 juillet 2010), en CD-ROM et URL : <http://www.linguistiquefrancaise.org/>

**Chesneau, A.** (2008) *Etude diachronique de deux corpus oraux : ESLO1 et ESLO2. Analyse contrastive de corpus par réenregistrement de locuteurs (1968-2008)*. Université d'Orléans, dir. G. Bergounioux.

**Eshkol I., Maurel D., Friburger N.** (2010), « Eslo : from transcription to speakers' personal information annotation », *Seventh language resources and evaluation conference (LREC 2010)*, Malte, 19-21 mai. *Actes en ligne* : <http://www.lrec-conf.org/lrec2010/>

**Guérin, E.** (2008). « Le français standard : une variété située », in Actes du I<sup>er</sup> Congrès Mondial de Linguistique Française 08, Paris, 9-12 juillet 2008.  
<http://www.linguistiquefrancaise.org/>

Gadet, F. & **Guérin, E.** eds (2008) *Etudes de syntaxe : français parlé, français hors de France, créoles*, Lixx 57.

**Serpollet, N.** (à par.) « How do you make an omelette? Study of the pragmatic and morpho-syntactic variations in the ESLO oral corpus », *Journal of French Language Studies*.

Tellier I., **Eshkol I.** , Taalab S., Prost J-P., (2010). « POS-tagging for Oral Texts with CRF and Category Decomposition", *Research in Computer Science, special issue : Natural Language Processing and its Applications* : 79-90

## LANGUES (Afrique, Amérique, Europe)

- Cristinoi, A.** (2006) « La Traduction Automatique comme terrain de recherche », *Translation Studies in the New Millenium. An International Journal of Translation and Interpreting* 4 : 67-74.
- Fabre, G.** (2009) « Sériation et hiérarchisation dans le système verbal du samba leko, langue Adamawa du nord-Cameroun », *Journal of African Languages and Linguistics* 30-2 : 223-266.
- Fournié-Chaboche, S.** (2009) « Les stéréotypes lexicalisés comme répétitions connotées dans une fugue littéraire : *Diálogo en Re Mayor* de Javier Tomeo », *Cahiers de Narratologie* 17  
URL : <http://revel.unice.fr/cnarra/index.html?id=1250>.
- Gallèpe, Th.** (2008) Zum Status der direkten Rede: Autonymie und Ikonizität. Wort und Text. Lexikologische und textsyntaktische Studien im Deutschen und Französischen. Festschrift für René Métrich zum 60. Geburtstag., Baudot, D. & Kauffer, M. (eds), Tübingen: Stauffenburg Verlag : 309 - 320.
- Grenand, F.** (2009) « Des Encyclopédies », in Grenand, F. (éd. en chef) *Encyclopédies palikur, wayana, wayâpi : langue, milieu et histoire*, fascicule 0, Orléans-Paris, PUO-CTHS : 10-34.
- Iglesias Ovejero, A.** (2006) « Interferencias lingüísticas y culturales en El Rebollar: algunas propuestas para la transcripción gráfica de conversaciones, relatos y testimonios orales en la modalidad cebollana » in *Cahiers du PROHEMIO* 7, *Interferencias lingüísticas y culturales en El Rebollar y otras partes*. Actas de las II Jornadas Internacionales de Estudio de El Rebollar (VI Coloquio PROHEMIO, 22-25 juillet 2005), P.U.O. : 43-73.
- Perrot, M.-E.** (2007) « Nature et fonction des stéréotypes linguistiques dans le discours de presse d'une communauté minoritaire : l'Évangéline de 1887 à 1930 », in Actes du colloque *Stéréotypage, stéréotypes : fonctionnements ordinaires et mises en scène* (Montpellier, juin 2006), Paris, L'Harmattan .
- Rougé J.-L.** (2008) « A inexistêcia de crioulo no Brasil », *Africa no Brasil*, J.-L. Fiorin & M. Petter eds., São Paulo, Editora contexto : 63-74.
- Serpellet, N.** (2008) « A Study of the Mandative Subjunctive in French and its Translations in English: A Corpus-Based Contrastive Analysis », *Incorporating Corpora – The Linguist and the Translator*, Anderman, G. & Rogers, M. (eds), Clevedon, Multilingual Matters : 187-216.
- Toupin, F.** (2010) « Exploring continuities and discontinuities between Ælfric's *Grammar* and its antique sources », *Neophilologus* volume 94, n° 2 : 333 sq.  
URL : <http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s11061-009-9185-5>

## MORPHOLOGIE ET PHONOLOGIE

**Abasq, V.** (2009) « Pre- et post- en anglais contemporain : sémantisme et comportement accentuel », *CORELA*, VII, 1.

URL : <http://edel.univ-poitiers.fr/corela/document.php?id=2115>

**Abasq, V. & Fournier, J.-M.** (2009) « The Phonology of English: Exceptions and Conflict », OCP6, *The Sixth Old World Conference in Phonology*, Edinburgh (AFF)

**Bergounioux, G.** (2006) CR d'*A Lateral Theory of Phonology* de T. Scheer, *Bulletin de la Société de Linguistique de Paris* CI, 2 : 112-127

**Bergounioux, G.**, Bergounioux M., Nguyen N. & Wauquier S. eds (2007) « Mathématiques et Phonologie » (avec M. Bergounioux, N. Nguyen et S. Wauquier), *Mathématiques et Sciences Humaines* 180, Paris, Editions de l'EHÉSS.

**Dugua, C.**, Spinelli, E., Chevrot, J.-P. & Fayol, M. (2009) « Usage-based account of the acquisition of liaison: evidence from sensitivity to plural/singular orientation of nouns », *Journal of Experimental Child Psychology* 102 (3) : 342-350.

**Fournier, J.-M.** (2007) « From a Latin syllable-driven stress system to a Romance vs Germanic morphology-driven dynamics » in Carr, P. & Honeybone, P. (Eds) *English phonology*, Language Sciences 29 : 218–236

**Fournier, J.-M.** (2010) *Manuel d'anglais oral*, Paris : Ophrys

**Osu, S.N. & Clements, N.** (2009) « Les nasales en ikwere », *CORELA* 7, 1.

URL : <http://edel.univ-poitiers.fr/corela/>

**Rougé, J.-L. & Emmanuel Schang** (2006) « The origin of the liquid consonant in Saotomense Creole », *The Structure of Creole words*, Niemeyer-Verlag, Tübingen : 23-29.

## SEMANTIQUE ET ANALYSE DU DISCOURS

**Cadiot, P.** & Visetti Y.-M. (2006) *Motifs et proverbes, Essai de sémantique proverbiale*, Paris, Presses Universitaires de France.

**Cadiot, P.** (2008) « Pragmatics of prepositions. A study of the French connectives *du coup* et *pour le coup* », in D. Kurzon & S. Adler (eds.) : *Adpositions, Pragmatic, semantic and syntactic perspectives*, Amsterdam : John Benjamins : 115-132.

**Col, G.** & Victorri, V. (2007) « Comment formaliser en linguistique cognitive ? Opération de fenêtrage et calcul du sens temporel. ». *CORELA —Cognition, Discours, Contextes*, G. Achard-Bayle, G. & et Paveau A.-M.

URL : <http://edel.univ-poitiers.fr/corela/document.php?id=1598>

**Col, G.** (2007) « Windowing the future. The cognitive operation of ‘windowing’ in the study of future time evocation », in *Du fait grammatical au fait cognitif / From Gram to Mind: Grammar as Cognition*, J.-R. Lapaire, et al. (eds), Bordeaux : Presses Universitaires : 323-341.

**Col, G., Osu, S., Garric, N. & Toupin, F.** (2010). « L'identité en construction ou de l'identité à la modalité » in *Construction d'identité et processus d'identification*, S. Osu, N. Garric, F. Toupin et G. Col (eds), Peter Lang. 1-12.

**Garric N.** & Léglise I. (2008) « Le discours patronal, un exemple de discours économique », *Mots* 86 : 67-84.

**Hamma, B.** (2007) « Philosophie et linguistique à la croisée des chemins ? La problématique de l'indicible », *Res Per Nomen*, P. Frath, Ch. Glendhill et J. Pauchard (Dir.), Reims, Epure, Reims : 195-212.

**Nemo, F.** (2006) « The Pragmatics of Common Ground. From Common knowledge to Shared Attention and Social Referencing? ». *Lexical Markers of Common Grounds* (Studies in Pragmatics, Volume 3), Fetzer, A. & Fischer, K. eds, Amsterdam, Elsevier Science : 143-158.

**Nemo, F.** (2007). « Reconsidering the Discourse Marking Hypothesis », *Connectives As Discourse Landmarks (Pragmatics and Beyond New Series)* Celle, A. & Huart R. eds, Amsterdam/Philadelphia, John Benjamins Publishing : 195-210.

**Toupin, F.** (2006) « Le Sémantisme primitif de *come* et de *go* », *L'Articulation langue-littérature dans les textes médiévaux anglais 3*, Presses de l'Université de Nancy : 73-97.

## DIDACTIQUE

**Bourdier, Ph.** (2008) Un grand écran pour les Lettres : le cinéma et l'enseignement du français, préf. C. Murcia, col. « Nouvelles pédagogies », éd. L'Harmattan, Paris, 313 p.

**Bourdier, Ph.** (2009) « Le cinéma en français : une question de représentation », *Le Français Aujourd'hui*, numéro : « Film et texte : une didactique à inventer », 165 : 51-63.

Gadet, F. & **Guérin, E.** (2008) « Le traitement du couple oral/écrit par la sociolinguistique et la description de la langue à visée didactique », *Le Français aujourd'hui*, 162 : 21-28.

**Lafont-Terranova, J. & Colin, D.** (2006) « Les enseignants de collège et l'écriture : des représentations à la formation », in Barré De Miniac, C. & Reuter, Y. (dir.), *Apprendre à écrire au collège dans les différentes disciplines*, Lyon, INRP : 59-84.

**Lafont-Terranova, J.** (2008) « Atelier d'écriture et compétence scripturale », *Enjeux* 71 : 87-121, Namur, CEDOCEF / Presses universitaires de Namur.

**Lafont-Terranova, J.** (2009) *Se construire, à l'école, comme sujet-écrivain : l'apport des ateliers d'écriture*, Namur, CEDOCEF / Presses universitaires de Namur (Diptyque 15).

**Lafont-Terranova, J., Colin, D. & Isidore-Prigent, J.** (2010) « Etude du rapport à l'écriture d'enseignants du primaire (CM2) et du secondaire (6<sup>e</sup>) dans une perspective curriculaire », XI<sup>e</sup> colloque international de l'Association internationale pour la recherche en didactique du français (AIRDF), *Quelles progressions curriculaires en français ?* Liège 26-28 août 2010. Actes sur CD-ROM.

# EQUIPES, AXES, THEMATIQUES

## **Equipes**

ESLO

Créoles et Langues d'Afrique

PROHEMIO

Langues de Guyane

GORDF

*Construction du sens* (thématique transversale)

## **Axes**

Analyse de discours

Morphophonologie

LTI

Diachronie de l'anglais

## **Actions**

CORELA

RSP

CAT-COD

RFP

VariLing

# ESLO

## Enquête Socio-Linguistique à Orléans

### Composition de l'équipe

Coordonnateur : Olivier Baude

Enseignants-chercheurs (EC) : L. Abouda, O. Baude, G. Bergounioux, G. Cloiseau, C. Dugua, I. Eshkol, E. Guérin, B. Hamma, M.-E. Perrot, N. Serpollet

Doctorants : L. Hriba, S. Moukrim

Collaborateurs : A. Chesneau, M. Perrin et une quinzaine d'étudiantes pour les transcriptions.

Autres EC du LLL associés à ESLO : J.-L. Rougé (programme LCO)

### Historique

L'équipe a été constituée progressivement à partir de 2000 à l'intérieur du CORAL avant qu'il ne devienne le LLL lorsque O. Baude, relançant un travail engagé dans les années 1990 autour du rapatriement de Colchester à Orléans des archives d'ESLO et d'une réflexion sur la constitution des corpus (cf. *Langue française* 93), a pris l'initiative de relancer l'exploitation du fonds et de préparer une nouvelle enquête qui permettrait d'établir une comparaison. Le programme, engagé sur fonds propres avec l'aide du conseil scientifique de l'université d'Orléans, a bénéficié à partir de 2007 d'un financement de l'ANR, projet corpus Variling. Ce programme, prolongé d'un an, arrive à échéance en janvier 2011.

### Etat de la recherche sur les corpus oraux

A la différence de ce que l'ATILF, reprenant les travaux de l'INaLF, a pu réaliser à Nancy sur l'écrit, ou du travail accompli au LDI, la recherche en France et sur le français pâtit d'une insuffisance de données orales disponibles fiables, facilement accessibles et indéfiniment extensibles. Il existe des corpus d'une grande qualité à l'étranger (Louvain et Leuven en Belgique, Montréal et Moncton au Canada) et dans les universités françaises (Aix, Lyon, Nice, Toulouse, Paris X, Paris 3, Paris XI...) cependant leur objectif a souvent une dominante thématique (phonétique, phonologique, linguistique de l'interaction, didactique...) et leur constitution correspond aux finalités de recherche d'un laboratoire ou d'un programme comme le montre le programme PFC (Phonologie du Français Contemporain). Les incitations du CNRS, à travers le CRDO et le TGE ADONIS, représentent des étapes importantes pour la structuration de ce domaine en émergence.

Parallèlement, il existe des demandes croissantes pour un corpus oral du français comparable à ceux qui sont développés pour d'autres langues. D'un côté, le marché très important du FLE est un consommateur important de données récentes et contrôlées, de documents authentiques destinés à l'enseignement audio de la langue. De l'autre, de nombreuses applications en Traitement Automatique des Langues (TAL) requièrent, pour leurs expérimentations, des données qui prennent en compte la variation des locuteurs et des situations. Un corpus qui permettrait d'échantillonner les expérimentations afin de rendre comparable les travaux effectués serait apprécié.

## Objectifs scientifiques : prototype et référence

Dans ce projet d'ensemble en cours de structuration, ESLO est centré sur la constitution d'un corpus de français oral qui réponde aux deux caractéristiques suivantes :

- avoir une valeur prototypique pour la constitution de corpus,
- pouvoir servir de référence pour les usages contemporains du français parlé.

**Concevoir un corpus prototypique** revient à caractériser l'ensemble de la chaîne de constitution des données comme une série d'opérations telle qu'elle permette, à chaque étape, de nouveaux développements et, dans l'enchaînement, qu'elle valide rétroactivement les reconfigurations nécessaires.

A minima, cette exigence conduit à définir explicitement les normes retenues, en les inscrivant dans les standards de la communauté (par exemple OLAC ou les Creative Commons), en veillant à conserver la comparabilité et l'interopérabilité. Concernant ce dernier point, une attention particulière est portée pour laisser le corpus ouvert à des interrogations qui ne peuvent encore être envisagées, notamment celles qui seraient issues d'autres disciplines que la linguistique (sociologie, psychologie, anthropologie, théorie du signal...).

Deux points ont fait l'objet d'une réflexion spécifique, l'un concernant l'objet, l'autre la procédure. Pour l'objet, on a tenté de réconcilier l'étude sociolinguistique qui porte sur une langue (et ESLO2, comme ESLO1, a suivi d'abord une approche de ce type) et celle qui étudie des interactions. ESLO2 a infléchi la collecte de données pour en faire un témoignage du français parlé tel qu'il est aujourd'hui en France, c'est-à-dire en relation avec les autres langues, dans les interactions, dans la pratique des locuteurs et dans les institutions : c'est dans cette perspective qu'a été conçu le programme LCO (Langues en Contact à Orléans).

Pour la procédure, les laboratoires se trouvent confrontés, dans la réalisation de corpus, à une activité qui requiert beaucoup de temps et d'argent. Toute amélioration, en particulier dans la transcription ou dans l'exploitation, qui réduirait ces deux formes d'investissement constitue un gain pour la communauté. Plusieurs avancées ont eu lieu dans ce sens :

- préparation des tâches par la rédaction d'un *Manuel du transcripteur et du relecteur*,
- analyse des erreurs (thèse de L. Hriba-Mcharek),
- méthodes d'entraînement à la transcription rapide,
- définition d'un cahier des charges pour l'exploitation d'un site de consultation ouvert de façon différenciée aux collaborateurs, aux chercheurs et au public (contrat avec la société Ares).

**Concevoir un corpus de référence** du français parlé aujourd'hui n'est pas un objectif que puisse s'assigner une équipe d'une dizaine de personnes. La définition des seuils quantitatifs (en millions de mots ou en heures transcrites) suppose une entreprise collective dès lors que sont pris en compte les critères de variation, la qualité des transcriptions, la protection juridique des témoins et des enquêteurs, la facilité d'accès aux données.

En revanche, et tel est l'objectif d'ESLO, il est possible de construire un corpus qui soit plus qu'une maquette et qui permette de mettre à l'épreuve, en grandeur réelle, les exigences d'un tel travail. Les ESLOs, sans ambition d'exhaustivité, ont néanmoins garanti une homogénéisation des données sur plusieurs centaines d'heures, ce qui en fait dès à présent le plus vaste corpus de français oral accessible.



Parallèlement, on a veillé à ce qu'il y ait dès à présent la consignation d'une variation sociale, dans la qualité des témoins, dans les thématiques abordées comme dans les formats d'échange. Avantage supplémentaire : ESLO est la seule enquête diachronique fiable puisqu'elle contraste les données recueillies en 1968 et celles d'aujourd'hui. Cette raison, qui donne un prix particulier au renouvellement sur le même terrain de l'enquête, conforte celles qu'avaient avancées nos prédécesseurs sur le caractère dialectologiquement neutre d'Orléans ou sur ses qualités urbaines sans pour autant focaliser l'étude sur Paris.

L'expérience acquise doit à présent s'intégrer dans un réseau dont les éléments institutionnels sont en place, qu'il s'agisse du CRDO et du TGE-ADONIS, des initiatives de l'ILF et surtout des nombreux travaux conduits dans les laboratoires, en France et à l'étranger.

## **Transcriptions et formats**

La linguistique de corpus, qui a connu un développement très important ces vingt dernières années, reste pour une grande part conditionnée par le modèle des corpus écrits. Le défi que représente la constitution de vastes corpus oraux et les méthodes optimales d'exploitation ont conduit à une double réflexion. D'un côté, il a fallu expérimenter les différents outils à disposition de la communauté afin de les évaluer et de les adapter, de l'autre une partie des connaissances a pu être validée dans des recherches comme celles menées par le LIMSI. Il est significatif qu'une entreprise comme Vecsys ait choisi de recruter des étudiants issus de nos formations pour assurer le développement de son département corpus oraux. D'autres collaborations ont été conduites avec le Laboratoire d'Informatique de Tours en particulier autour des disfluences.

Les transcriptions ont fait l'objet d'une attention particulière. Il ne s'agissait pas seulement de disposer du maximum d'heures de transcriptions alignées et relues mais de réfléchir à chacune des étapes sur les opérations, comme il avait fait en amont sur le choix des matériels, la définition des conditions d'enregistrement, la sélection des modules et des témoins. La tâche assignée à l'équipe, dans ce cadre, présentait l'avantage qu'elle recoupait les demandes formulées par d'autres confrontées à la notation de langues peu ou pas transcrites en Afrique et en Amérique du sud. La confection d'un guide de transcription a prolongé les indications apportées par le *Guide des bonnes pratiques*.

La pratique de la transcription a conduit les chercheurs à développer, à côté du travail ordinaire de collecte, d'archivage, de traitement et d'analyse, quatre points en particulier :

1. les méthodes de codage et de catalogage afin de rendre accessibles les données, en particulier en assurant la mise en ligne : une journée d'étude avec des anthropologues à l'initiative de R. Walter et un colloque organisé à Orléans par E. Schang ont permis de faire le point sur les problèmes ;
2. l'interprétation des erreurs d'écoute et la multiplicité des rédactions obtenues par différents transcripateurs dans le rapprochement des notations effectuées sur les mêmes enregistrements : la thèse de L. Hriba traite spécifiquement de ce point à partir d'outils logiciels dédiés et trouve un prolongement dans certains travaux du GORDF,
3. la question de l'anonymisation afin de pouvoir concevoir une mise en ligne qui protège les personnes ayant accepté de répondre à l'enquête tout en préservant au maximum l'intégrité des contenus disponibles,
4. la définition d'un site organisé autour des corpus oraux en partenariat avec la société Ares avec une différenciation des requêtes entre le grand public, la communauté savante et les chercheurs de l'équipe.

L'élévation du niveau d'exigence (conserver tous les états de transcription, préserver le signal et assurer l'alignement sur le texte, recourir exclusivement à des logiciels facilement disponibles...) a rendu plus complexe, et donc plus délicate, la réalisation mais elle a constitué un banc d'essai irremplaçable. L'expérience accumulée à cette occasion semble directement exploitable pour un projet d'ensemble.

A noter qu'en collaboration avec le GORDF, un module est en cours de développement en milieu scolaire qui concerne aussi bien la collecte de données auprès d'enfants, avec des questionnaires adaptés que la mesure d'une diachronie en temps réel par la comparaison d'enquêtes conduites simultanément auprès des enfants et de leurs parents afin de mesurer l'évolution, dans un même milieu, des productions verbales. Sur un plan interlinguistique, une question comparable se pose avec l'étude de la « linguistic loyalty » dans les communautés allophones.

### **Langues en Contact à Orléans (LCO)**

Ce module a bénéficié de soutiens du Conseil Régional du Centre et de la Délégation Générale à la Langue Française et aux Langues de France (Ministère de la Culture) et de l'Université d'Orléans.

Force est de constater qu'il n'existe actuellement aucune monographie qui, à l'échelle d'une agglomération française, recenserait l'ensemble des langues qui y sont parlées de façon continue, quelle qu'en soit la raison. Combien aujourd'hui parle-t-on de langues à Orléans (250.000 Habitants) ? Il est sûr qu'il y en a plus de 200 mais encore ? En dénombre-t-on 400 ? Ou même 600 ? Plus encore ? Ce sont de tels ordres de grandeur auxquels on est confronté, en même temps que se repose la question de savoir quel est le statut qu'il convient de donner à ce que découvre l'enquête : est-ce une langue ou un dialecte ? A partir de quel degré de maîtrise doit-on valider sa présence ? Par combien de personnes ? Que faire des « mélanges de langues » ? A supposer qu'il se rencontre au moins un locuteur d'une langue donnée dans l'agglomération d'Orléans, comment la langue est-elle pratiquée (ou non) au quotidien.

On soulignera dans le bilan combien sont difficiles de telles études dans le contexte français, et pas seulement si l'on en juge par une expérience comparable menée à Barcelone. La « linguistic loyalty » n'est pas affichée comme telle et la rencontre avec les locuteurs, à une telle échelle, requerrait des moyens comparables à ceux dont peut disposer l'INSEE qui ne consent pas à livrer ses résultats à une telle échelle.

Pour engager la recension, des études ont été conduites en milieu scolaire et universitaire, en 2007-2008 dans une école de la ville de Saint Jean de la Ruelle en collaboration avec le CASNAV, l'Inspection de l'Education Nationale, la direction de l'école et la municipalité. En 2008-2009, l'enquête a été élargie à toutes les écoles abritant une Classe d'Initiation et à tous les collèges d'Orléans. Dans chaque établissement, les élèves de deux classes (CE2/CM2 et 5<sup>e</sup>/3<sup>e</sup>) ont été interviewés. En parallèle, deux enquêtes avec un questionnaire du même type ont été réalisées, l'une à la sortie des lycées, l'autre sur le campus de l'Université.

Parallèlement, une série d'entretiens semi-directifs avec des interprètes ou des personnes ressources de diverses communautés linguistiques (arabe, berbère, lingala, mandinka, serbo-croate, tamoul, vietnamien etc.) a été menée. Ils portent sur la vie des différentes langues à Orléans et les parcours migratoires. Ils fournissent des échantillons de différentes façons de parler français en situation de contact de langues.

Mentionnons encore un relevé photographique d'enseignes de magasin en langues étrangères réalisé en 2008-2009. L'ensemble de ces travaux est conduit en intégrant les étudiants des modules de sociolinguistique en licence et Master. Une thèse de doctorat sur les représentations du plurilinguisme en milieu scolaire s'inscrit entièrement dans le cadre du projet LCO. Trois autres (sur l'arabe et le berbère, le peul, le capverdien) ont, entre autres terrains d'applications, la ville d'Orléans et rejoignent ainsi les préoccupations de ce programme.

## **Formation**

La recherche conduite dans le cadre d'ESLO a eu un impact sur l'enseignement en sciences du langage. Dès la licence, un certain nombre d'étudiants ont été associés à des fragments de l'enquête et une douzaine d'entre eux a été recrutée, sur des contrats étudiants, pour effectuer transcriptions et relectures, encadrée par des doctorants ou des enseignants. Ces contrats ont été prolongés en Master pour quelques-uns.

Quatre thèses ont été soutenues (ou sont en cours de rédaction finale) à partir des corpus ESLO :

A. Chesneau sur la comparaison de dix locuteurs d'ESLO1 repris dans ESLO2

M. Petit sur l'étude prosodique des connecteurs

S. Moukrim sur l'expression temporelle en français, berbère et arabe marocain

L. Hriba sur la multi-transcription.

Un transfert de connaissances a été opéré lors d'une mission au Burkina sur financement de l'AUF afin de présenter aux collègues du CNRST et de l'Université les méthodes de traitement des corpus oraux qui sont une des ressources linguistiques essentielles de la phonothèque nationale de ce pays.

## **Vulgarisation**

En terme de vulgarisation, on notera quelques exposés grand public et des articles dans *Langues & Cités* et dans *Pour la science* (M. Petit).

# Créoles et langues d'Afrique

## Composition de l'équipe

Coordonnateur : Jean-Louis Rougé

Enseignants-chercheurs (EC) : A. Delplanque (e. r.), G. Fabre, S. Osu, J.-L. Rougé, E. Schang

Doctorants : M. Cavalheiro, I. Diallo, D. Evora, B. Magnana, E. Traore

Collaborateurs : R. Boyd

Laboratoires français impliqués : CEDETE, IRD, LLACAN, LLF, SFL

Coopération internationale : Université de Coimbra (Portugal), Université de Sao Paulo (Brésil), centre d'enseignement supérieur de Praia (Cap-Vert), CNRST et université de Ouagadougou (Burkina).

## Historique

La constitution de cette équipe remonte à 2003 quand deux chercheurs travaillant sur les créoles portugais d'Afrique dans d'autres cadres académiques ou professionnels se sont retrouvés réunis, suite à une politique de recrutement volontariste, à l'Université d'Orléans. Le recrutement de G. Fabre, dont la thèse porte sur le samba-leko (Cameroun) puis la fusion avec Tours où A. Delplanque est le spécialiste international du dagara (Burkina) et S. Osu de l'ikwéré (Nigeria) ont permis d'étendre à d'autres langues d'Afrique sub-saharienne les investigations. Les travaux sur le Cameroun et le Nigeria ne sont pas repris ici car ils ont été conduits avant l'accueil de ces chercheurs dans le laboratoire.

## Etat de la recherche sur les créoles portugais d'Afrique

### *Constitution de corpus sonores*

La constitution de corpus sonore des créoles portugais d'Afrique et de variétés émergentes dans les contacts luso-africains répond à un double objectif : mettre à disposition des documents permettant l'étude de ces langues ; conserver les traces de langues pour celles qui sont en danger de disparition et pour les variétés linguistiques éphémères. Actuellement, le groupe possède des enregistrements des créoles de Guinée Bissau et de Casamance, du Cap-Vert, des différents créoles de São Tomé ainsi que du portugais des Tonga.

Au cours des quatre dernières, deux enquêtes de terrain ont été effectuées essentiellement en Guinée Bissau et en Casamance (Sénégal), une région dans laquelle l'instabilité politique et sociale provoque des déplacements de population dont les conséquences sur la vie des créoles sont sensibles. Des enregistrements de témoins d'origine sociale et ethnique et d'âge différents ont été collectés. Le travail de transcription et d'exploitation des données se poursuit. La diversité du matériel recueilli a permis une approche comparative de ces créoles qui a fait l'objet de présentation dans différents colloques (voir publications).

### *Coopération nationale et internationale*

Les spécialistes de créoles du LLL participent activement à l'organisation du séminaire mensuel

sur les grammaires créoles organisé à Paris-8 (UMR 7023) et qui réunit des chercheurs français et hollandais. Une journée d'étude internationale sur les grammaires créoles, en collaboration LLL, SFL, LLF, a été organisée le 25 juin 2010 à Orléans. E. Schang participe au projet Creoloral de la fédération TUL en partenariat avec la DGLFLF et il apporte son expertise pour le recueil et l'exploitation des données orales sur les créoles français de la Guadeloupe et de la Martinique. Il participe également au projet franco-hollandais sur la réduplication dans les langues créoles (coord. Enoch Aboh, UvA). Un projet d'assistance à Haïti, en particulier au Centre de Linguistique Appliquée de l'Université d'Etat, est actuellement en cours à l'initiative de l'Université d'Orléans.

Les chercheurs de ce programme sont impliqués depuis plusieurs années dans l'Association des Créoles de Base Lexicale Portugaise et Espagnole (ACBLPE) et ont occupé ou occupent diverses responsabilités (secrétariat, présidence de l'Assemblée générale, etc.) au sein de cette association. Actuellement, le président de l'association est J.-L. Rougé.

Au nombre des réalisations conduites dans ce cadre, on mentionnera l'organisation de colloques annuels (après Orléans, les derniers ont eu lieu à Amsterdam, Lisbonne et Cologne avant Villejuif en 2010 à l'initiative du LLACAN) et la mise en place du site Internet de l'association qui est domicilié à Orléans.

L'ACBLPE a permis de tisser des liens spécifiques avec des universités étrangères. Actuellement, un doctorant de l'Université de Coimbra est en stage à Orléans où il participe aux travaux de l'équipe. De façon plus ambitieuse, l'équipe contribue à la mise en place de formations universitaires dans les pays de l'Afrique lusophone, en particulier au Cap-Vert.

### **Etat de la recherche sur les langues du Burkina**

Le LLL bénéficie, dans le travail accompli au Burkina, de l'expérience d'Alain Delplanque qui, en retraite depuis trois ans, a poursuivi les recherches entreprises sur le terrain lors de ses années d'enseignement dans ce pays. Spécialiste du dagara sur lequel il a fait une thèse remarquée, il a depuis poursuivi ses investigations et fait encore profiter l'équipe de son expérience et de ses contacts.

Depuis une dizaine d'années, une collaboration, d'abord assurée par un financement de l'ANR, a été par ailleurs établie entre le laboratoire de linguistique de l'université d'Orléans et le département linguistique du CNRST du Burkina. Une première étape du travail a consisté dans un échange d'expériences comprenant des missions conjointes sur le terrain, en France et au Burkina. Une mission d'enseignement pour un collègue de Ouagadougou, L. Sanogo, aujourd'hui conseiller technique pour les langues nationales auprès du Ministère chargé de l'alphabétisation et de l'éducation non formelle, a permis une confrontation des méthodes. Une session de formation à la collecte et au traitement des corpus oraux organisée sur place a permis d'assurer un transfert des techniques et une diffusion des compétences.

Depuis un an, c'est un nouveau projet qui a été impulsé, grâce à un financement de la Région Centre : RADICEL-K, en collaboration avec l'IRD. Même si le décalage calendaire fait qu'en janvier 2012 ce projet aura déjà largement réalisé les études auxquelles se sont engagés les collègues, on a choisi de le faire figurer dans le programme plutôt que dans le bilan. En effet, l'essentiel du travail reste à accomplir à l'heure où ces lignes sont rédigées et la collaboration des

partenaires autour de la description des langues rejoint largement la présentation sur les corpus oraux qui constituent le cœur de l'action entreprise par le LLL dans le prochain contrat.

### **Formation doctorale**

Ces programmes alimentent l'organisation de différents séminaires de Master. Ils fournissent les thèmes d'étude de cinq thèses en cours de rédaction dont une devrait être soutenue d'ici la fin de l'année.

4<sup>e</sup> année

M. Cavalheiro (allocation de l'université d'Orléans puis ATER) : le français au Burkina

2<sup>e</sup> année

E. Traore (ingénieur d'études au CNRST) : Le tagba (Burkina Faso)

I. Diallo (chargé de recherche au CNRST) : Le système verbal du peul au Burkina Faso

D. Evora (BGF) : Compétences langagières et parcours migratoires dans la diaspora capverdienne

Entrée en thèse en septembre 2010

B. Magnana (contrat doctoral) : description d'un dialecte kota

# PROHEMIO

Programme de Recherche sur « Oralité, Histoire, Ecriture » dans le Monde Ibérique, d'Orléans

*Responsable* : S. Fournié et F. Iglesias Ovejero (professeur émérite)

*Doctorants* : C. Hounnouvi, Y. Haquin,

*Autres participants* : J. Fernandez, F. Giraud

Le PROHEMIO a connu depuis 2006 une restructuration conséquente puisque, suite à la retraite de deux de ses membres, il est passé de 3 EC (un PU et deux MCF) à un seul (plus un PU émérite). Depuis deux ans, deux doctorants sont venus renforcer les effectifs qui se composent à l'heure actuelle de :

Depuis sa création en 1993 le PROHEMIO s'est fixé pour objectif de travailler sur l'oralité dans le monde ibérique, que ce soit en relation avec l'écriture à travers notamment l'étude des formes fixes, ou qu'il s'agisse d'étudier les rapports entre oralité, histoire et culture traditionnelle. Les travaux actuels du PROHEMIO s'articulent donc autour de ces deux axes.

## **Oralité et écriture : étude du figement linguistique**

Le PROHEMIO développe une analyse linguistique contrastive fondée sur l'étude du figement et la traduction avec pour projet l'édition d'un *Dictionnaire idéologique comparatif des locutions espagnoles et françaises*

Après la réorganisation de l'équipe, le projet initial de dictionnaire historique des expressions figées a été remodelé en l'adaptant aux contraintes actuelles du groupe et à l'évolution de la lexicographie.

-Une phase de réflexion a permis d'élaborer un projet neuf en introduisant une dimension idéologique (on commencera par un dictionnaire des défauts) et un support informatique (réalisation du dictionnaire à l'aide du logiciel Toolbox). Cette technique doit permettre de publier des éléments du dictionnaire, au fur à mesure de sa constitution, en l'hébergeant sur un site dédié.

-Une phase de travail préparatoire a abouti à la constitution d'un corpus répertoriant toutes les expressions figées relevées dans différents romans d'auteurs espagnols contemporains. Une recherche croisée portant à la fois sur des textes espagnols non traduits (l'intégralité des écrits de Javier Tomeo, une quarantaine de titres, a été analysée) et sur des textes littéraires espagnols traduits en français. Dans ce cas, le relevé porte à la fois sur le texte espagnol et la traduction. A ce jour, des romans de Cela, Delibes, Marsé, Martín Gaité, Mendicutti, Mendoza, et Tomeo ont été analysés (cf. Annexe : "Corpus actuel d'ouvrages traduits").

- La réflexion porte actuellement sur l'agencement du dictionnaire. Il s'agit en effet de constituer un outil d'aide à la traduction utile aux chercheurs, aux étudiants et aux traducteurs. Il a pour objet de mettre en parallèle le trésor lexicographique des deux langues tout en respectant « les cases vides », c'est-à-dire l'absence de locution ayant tel sens dans l'une des deux langues. Au plan méthodologique, le dictionnaire est classé de façon thématique. A l'intérieur d'un thème sont distinguées les principales nuances liées aux différentes acceptions et les types d'images auxquelles elles renvoient. Les essais de présentation et d'utilisation du support informatique sont en cours. La thèse de Y. Haquin porte précisément sur cet aspect du travail.

Le PROHEMIO s'intéresse également aux relations entre oralité et discours littéraire. Dans ce cadre, plusieurs articles ont été publiés depuis la création de l'équipe, notamment en lien, pour les plus récents, avec la stéréotypie, la répétition et la réécriture.

## Les Colloques PROHEMIO

Une activité centrale du PROHEMIO est, depuis sa fondation, l'organisation de colloques qui, partant des formes de l'oralité présentes dans les textes, en particulier littéraires, est allée progressivement en se spécialisant dans les récits de vie et le témoignage historique oral. Une attention particulière aux documents collectés autour de la guerre civile ont alimenté les réflexions des chercheurs.

20-22 juillet 2007

*Memoria histórica – Historia oral : guerra, postguerra y emigración en El Rebollar y otras partes*, VII<sup>e</sup> Colloque PROHEMIO organisé en collaboration avec l'Association *Documentación y Estudio de El Rebollar* et l'Université de Salamanque à Ciudad Rodrigo, Robleda et Navasfrías.

17-19 juillet 2009

*Historia, habla y cultura tradicional*. VIII<sup>e</sup> Colloque PROHEMIO, organisé en collaboration avec l'Association *Documentación y Estudio de El Rebollar*, à Fuenteguinaldo, Robleda et Ciudad Rodrigo (Salamanque).

Parallèlement, sont parus les actes :

Iglesias Ovejero, A. dir. (2006) *Cahiers du P.R.O.H.E.M.I.O.*, n° 7 et n° 8, *Interferencias lingüísticas y culturales en El Rebollar y otras partes*. Actas de las II Jornadas Internacionales de Estudio de El Rebollar (VI Coloquio P.R.O.H.E.M.I.O.), 22-25 juillet 2005, publiés avec la collaboration de Lidio J. Fernandez et Françoise Giraud. P.U.O., 315 pp. et 451 pp.

Iglesias Ovejero, A., dir. (2008) *Cahiers du P.R.O.H.E.M.I.O.* n° 9 et n° 10, *Memoria histórica – Historia oral : guerra, postguerra y emigración en El Rebollar y otras partes*. Actas de las III Jornadas Internacionales de Estudio de El Rebollar (VII Colloque P.R.O.H.E.M.I.O.), 20-22 juillet 2007, publiés avec la collaboration de Lidio J. Fernandez, Françoise Giraud et Cécile Iglesias. P.U.O., 406 pp. et 458 pp.

Iglesias Ovejero, A. dir. (2010) *Cahiers du P.R.O.H.E.M.I.O.*, n° 11, DVD, *Historia, habla y cultura tradicional*. Actas de las IV Jornadas Internacionales de Estudio de El Rebollar (VIII Coloquio P.R.O.H.E.M.I.O.), 17-19 juillet 2009, publiés par A. Iglesias Ovejero avec la collaboration de Françoise Giraud et Beatriz Sánchez Valdelvira. P.U.O. Presses Universitaires d'Orléans, 2010, 560 pp.

Avec des moyens limités, le PROHEMIO a poursuivi une double activité scientifique, dans la lexicographie des figements et dans l'étude de l'oralité. Il a prolongé ses travaux par une application au « naming » et à l'analyse sémiotique du figement linguistique dans la communication d'entreprise. Différents articles témoignent de cet intérêt ainsi que l'inscription de la thèse de C. Hounnoui : *Pour une approche interculturelle de la communication marketing dans le monde hispanophone*.



# Langues de Guyane

## Composition de l'équipe

Coordonnateur : F. Grenand

Enseignants-chercheurs (EC) : P. Cadiot, A. Cristinoi, F.Nemo

Collaborateurs : P. Grenand et les collaborateurs de l'*Encyclopédie*

Coopération internationale : Université de Sao Paulo (Brésil)

La direction de programme a été assurée par Françoise Grenand, DR CNRS en détachement à Cayenne pour assurer la direction du programme Oyapock. Elle est assistée d'Antonia Cristinoi, Pierre Cadiot et François Nemo pour la partie linguistique et de Pierre Grenand (DR IRD) pour la partie anthropologique. Le programme remonte, dans ses premières missions, à 1998. Depuis, il s'est poursuivi sans interruption comme en attestent les publications et les travaux.

## Description scientifique de trois langues / trois cultures : l'*Encyclopédie*

L'objectif final des travaux entrepris par le laboratoire demeure la description exhaustive de trois langues, le palikur, le wayana et le wayampi, l'accent étant mis prioritairement sur la première de ces trois langues.

Pour chacune d'elles, il s'agissait d'établir les outils suivants, en faisant appel, autant que de besoin, à des collaborations extérieures. D'abord, une description phonologique et une application à la transcription. La disponibilité et l'accessibilité des documents recueillis par les chercheurs en Guyane supposent au préalable une transcription qui assure une consignation efficace des relevés effectués auprès des témoins. Ensuite, un dictionnaire, une grammaire descriptive et une encyclopédie (papier et électronique) qui prend corps autour des grandes sections suivantes :

- la présentation générale,
- les études thématiques,
- la bibliothèque,
- le dictionnaire,
- les index.

L'encyclopédie se présente comme une collection de publications autonomes à visée globalisante. Cela présente au moins deux avantages importants : sans dépendre les uns des autres, les auteurs publient leurs études à leur rythme, au fur et à mesure de l'achèvement de leurs travaux ; quant aux coûts, répartis par fascicule dans le temps, ils sont plus faciles à assumer.

Qu'ils soient ethnologues ou linguistes, les chercheurs n'ont jusqu'à présent guère été tentés par cette forme de restitution, laquelle répond pourtant à une forte demande sociale. En effet, au terme d'une longue pratique du terrain et des hommes, il ressort que, conscientes des dangers que représente l'intrusion, souvent brutale, de la modernité dans leur genre de vie, les sociétés ayant fondé leur mémoire ainsi que la transmission de leurs savoirs et savoir-faire sur une culture orale, croient aux vertus de la chose écrite.

Les encyclopédies, plus que la description de mots cherche à dresser le portrait du monde par un peuple. Les catégories du vivant, la conception du surnaturel, l'ordonnement de l'univers, le déroulement des grandes étapes de la vie, les arts et les techniques, sont autant de macro-thèmes dont l'approche se fait du dedans, avec pour fil d'Ariane la langue, dans son fonctionnement comme dans ses réalisations. Les ouvrages allient la langue vernaculaire en tant que langue source, le français en tant que langue cible et métalangage, les langues de contact en tant que langues de traduction spécialisée. (voir annexe Encyclopédie des Peuples de Guyane)

## **Les dictionnaires**

Une première version du dictionnaire du wayãpi a été réalisée par F. Grenand en 1986. Comportant 6000 entrées et 558 pages, il s'agit de le mettre à jour : il existe aujourd'hui plus de 3000 addenda ou correctifs à quoi s'ajoute un lexique de la modernité rassemblé au fil des années. Ont été financés :

- une mission centrée sur le recueil et la transcription des corpus oraux,
- la numérisation des données sonores des années soixante-dix et de leur transcription,
- le report des occurrences rencontrées en corpus dans un corpus d'emplois,
- la saisie numérique du dictionnaire existant.

Dans le même esprit, une base de données lexicales du palikur (arawak) est en cours, fondée sur la transcription et la traduction de 11 corpus oraux recueillis lors de trois missions sur le terrain effectuées entre 2004 et 2009. Ces corpus, qui doivent être intégrés dans un atlas sonore des langues de Guyane et s'intégrer dans les archives du laboratoire en dépôt à la BnF et auprès d'ADONIS, doivent aboutir à la confection d'un dictionnaire fondamental bilingue français/palikur et palikur/français, préparé en lien avec l'édition de l'*Encyclopédie*.

Une attention particulière sera portée à la question de l'expression du nombre en palikur et aux différents aspects de la classification nominale, dans une perspective typologique et contrastive.

## **Un projet monographique : l'alternance codique en contexte multilingue**

Les Palikur sont l'un des six groupes amérindiens de Guyane française. Présents aussi au Brésil, ils sont installés sur l'Oyapock et ses affluents ainsi qu'à Macouria. Le brassage des différents groupes et les alliances matrimoniales créent un contexte multilingue à quatre langues (palikur, créole, français, brésilien) avec une exposition constante aux unes et aux autres en fonction des moments et des périodes de la vie (par exemple, des enfants suivant en même temps des programmes en français et en portugais du Brésil sur deux télévisions et discutant en palikur). Le quadrilinguisme est celui d'une société : la maîtrise des langues (en particulier du français et du créole) étant très inégale.

Ce projet entend porter un diagnostic global sur la réalité du multilinguisme sous la forme d'une enquête testant les capacités de production et de compréhension d'un échantillon de locuteurs dans chacune des quatre langues, à l'oral et, pour les locuteurs alphabétisés, à l'écrit.

# GORDF

Groupe Orléanais de Recherche en Didactique du Français

## Composition de l'équipe

Coordonnateur : Jacqueline Lafont

Enseignants-chercheurs (EC) : Ph. Bourdier, E. Guérin, F. Wolf

Doctorants : C. Brumelot, C. Sarré (inscription principale au Havre), D. Schwob

Collaborateurs : D. Colin, A.-L. Doyen, G. Guetemme, J. Isidore-Prigent, A. Vaslin-Chesneau, D. Ulma et des collaborateurs de l'IUFM et de l'enseignement du premier et du second degré.

Coopération internationale : Université Catholique de Louvain (Belgique), Université de Laval (Canada), Université de Sherbrooke (Canada).

## Objectifs et collaborations

Le Gordf entend appliquer les méthodes et les outils de la linguistique aux usages scolaires et universitaires du français. Les recherches développées par les membres articulent, autour de problématiques centrées sur la didactique du français langue maternelle et plus particulièrement sur la didactique de l'écrit des questions concernant l'acquisition et l'évaluation avec une analyse des productions dans les différents niveaux de la formation initiale ou continue. Plus largement, le groupe intègre la didactique de l'image et la didactique des langues.

A la suite d'une journée d'étude sur la didactique de l'écrit (13 mai 2005) qui a permis à des chercheurs francophones (France, Québec, Belgique) de débattre sur la question de *La construction des savoirs et le sujet-écrivain*, le groupe a développé les collaborations nationales et internationales. Des collaborations ont ainsi été développées autour de la didactique de l'écrit avec la Belgique (UCL : Francine Thyron et Silvia Lucchini professeures, Maurice Niwese docteur ; CEDOCEF, FUNDP Namur) et le Québec (CRIFPE-Laval : Suzanne-G. Chartrand professeure agrégée à l'Université Laval et Christiane Blaser, professeure adjointe, Université de Sherbrooke). En outre, deux membres du GORDF participent au GDR CNRS 2657 qui réunit des chercheurs nationaux de disciplines variées autour de la *Production Verbale Ecrite*. Enfin, depuis 2007, le GORDF est impliqué dans le Plan Pluri-Formations *Ecole, Education et Sociabilités : Les conditions de maîtrise de la production d'écrits à l'école* piloté par l'université de Poitiers.

Les activités de recherche du groupe donnent lieu à des communications dans des colloques nationaux et internationaux et à la publication d'articles et d'ouvrages de didactique en France et à l'étranger. Un colloque national (*Images à l'école, image de l'école, 1880-1960*, Chartres, 17 mars 2010) a été organisé. L'organisation d'un séminaire régulier ouvert aux chercheurs, aux praticiens et aux étudiants (cf. annexe) permet au GORDF d'être non seulement un lieu de recherche mais aussi un lieu d'échanges et de formation, ce que confirme son implication dans l'encadrement de doctorants, la formation des étudiants et celle de formateurs de l'IUFM Centre Val de Loire

## Recherches

### *Didactique de l'écrit*

La didactique de l'écriture est principalement concernée via la participation à un PPF et des travaux qui ont permis des échanges avec la Belgique, également une implication dans les

réflexions du GDR CNRS 2657 *Production verbale écrite*. Il s'agit d'analyser les effets d'ateliers d'écriture créative inspirés du modèle de loisir et favorisant la réécriture, sur le développement de la compétence (méta)scripturale d'étudiants de deux filières universitaires (Lafont-Terranova 2007, 2008b, 2009 ; Lafont-Terranova & Niwese, 2010) : *atelier d'écriture et développement de la compétence scripturale dans une filière technologique* (2007-2009) et *atelier d'écriture et conceptualisation de savoirs didactiques sur la compétence scripturale dans un master de linguistique et didactique* (2008-2011). La didactique de la littérature est concernée via une recherche doctorale (cf. 4).

### ***Didactique de l'image***

Deux projets s'inscrivent dans ce thème : *Les interactions image-texte dans les albums de littérature*, projet IUFM CVL et *L'histoire du cinéma éducateur (1950-1960) : discours et pratiques d'enseignement de l'analyse des images cinématographiques en contexte scolaire*, projet de recherche commun aux universités d'Orléans et Tours) (Bourdier, 2010).

### ***Interactions orales en situation d'enseignement/apprentissage***

Ce thème a été abordé à travers trois axes de recherche de l'IUFM Centre Val de Loire : *Evaluation du langage oral en Grande Section de Maternelle* (2004-2007), *Les reformulations de l'enseignant en sciences et en français au cycle 3 de l'école primaire* (2005-2008) (Isidore-Prigent, 2007) et *Les interactions orales-écrites dans les situations d'apprentissage portant sur l'analyse d'images cinématographiques* (Bourdier, 2009).

### ***Didactique des langues***

Notons la participation à une réflexion sur le portfolio européen pour les enseignants de langue et sur son expérimentation (Meheut & Wolf-Mandroux 2009 ; *Recherche-action sur l'élaboration, l'utilisation et l'évaluation du portfolio des langues en primaire et début secondaire*, Université de Zagreb IUFM CVL, Wolf-Mandroux en cours), à deux ateliers du Centre Européen des Langues Vivantes (*Le pilotage et la mise en œuvre du Portfolio européen pour les enseignants en langues en formation initiale* 2008-2010 Wolf-Mandroux, en cours, & *L'enseignement précoce des langues par les contenus* 2009-2011) ainsi qu'une recherche doctorale (cf. 4).

### ***Acquisition***

Une étude longitudinale (sur 5 ans), soutenue par l'IUFM Centre Val de Loire, avait pour objet de mesurer les effets à long terme d'un entraînement phonologique à partir de la classe de moyenne section (Bianco, Bressoux, Doyen, Lambert, Lima, Pellencq & Zorman, 2010).

## **Participation au PPF Ecole, Education et Sociabilités :**

Ce PPF traite des conditions de maîtrise de la production d'écrits à l'école.

- axe socio-didactique : La question du rapport à l'écriture au moment du passage CM2 / 6<sup>e</sup>. Il s'agit de déterminer les conditions d'une évolution nécessaire du rapport à l'écriture des élèves et de la conception par les enseignants de situations didactiques adaptées. Les données orales (entretiens semi-directifs) et écrites (questionnaires) concernent une centaine d'élèves, 19 enseignants de primaire et de collège (français et SVT), (Lafont-Terranova, Colin & Isidore-Prigent, 2010). Une réflexion est en cours sur l'adéquation entre les objectifs de la recherche et le degré d'« oralité » à conserver dans les transcriptions d'entretiens.
- axe psychologie cognitive : deux thèmes sont développés. 1/ Dans le cadre de la liaison Grande Section/CP, montrer l'influence de la langue maternelle dans les premières productions des élèves en écriture approchée (« invented spelling »), (Doyen & Lambert, 2010 acceptée). 2/ Dans le cadre de la liaison lycée/supérieur, montrer que les abréviations sont sous-utilisées par les élèves et étudiants dyslexiques-dysorthographiques et caractériser le rôle des compétences langagières qui en permettrait une utilisation efficace.

En lien avec le travail mené dans ce PPF, le GORDF a été impliqué dans le colloque international *De la France au Québec : l'écriture dans tous ses états* (12-15 novembre 2008, Poitiers) : participation au comité scientifique, animation de deux ateliers, communication (Doyen & Lambert, 2008).

## **Formation**

Encadrement de doctorants : codirection (didactique de l'anglais de spécialité, soutenance prévue en 2010), coencadrement au sein du LLL (variation linguistique et didactique de la littérature), participation au comité d'accompagnement et au jury d'une thèse (janvier 2010), UCL, Belgique (didactique de l'écriture).

Parcours *Didactique du français langue maternelle* du Master *Linguistique et didactique* de l'université d'Orléans : responsabilité de trois UE mises en place en 2006-2007 (*Séminaire de didactique du français* 9h CM et 9h TD, *Situations didactiques* 48h TD, *Ateliers d'écriture et compétence scripturale* 24h CM et 24h TD) et encadrement de mémoires de M2 en lien avec la recherche menée dans le cadre du PPF.

Participation au plan de Formation de formateurs, IUFM CVL (4 formations sur 5 dans l'UCFR français en 2009-10) : inscription au plan de formation de formateurs, des 3 UE du master *Linguistique et didactique* (parcours FLM), deux formations à la recherche  
- *Didactique de l'image et production d'albums narratifs – passage GS-CP*  
- *Liaison CM2-6ème - Le rapport à l'écriture - Travail en collaboration avec une équipe de recherche* : formation articulée avec le PPF.

Interventions à l'IUFM Poitou-Charentes, au CEDOCEF (FUNDP, Belgique).

Expérimentations au département informatique (IUT d'Orléans) en lien avec la recherche sur les ateliers.

## **Séminaire**

Le GORDF organise un séminaire régulier en faisant intervenir des chercheurs de l'université d'Orléans et d'autres universités françaises ou étrangères (Belgique, Québec). Le séminaire dont l'entrée est libre a trois fonctions : confrontation entre chercheurs et diffusion de la recherche, formation initiale (deux U.E. du master *Linguistique et Didactique* ont intégré ce séminaire) et formation de formateurs (le séminaire est inscrit au Plan de formation de formateurs de l'IUFM CVL depuis 2007-2008). Depuis 2007-2008, 23 conférences (cf. annexe *Séminaire GORDF 2007-2008 à 2009-2010*) ont été organisées, permettant des rencontres entre chercheurs, formateurs, praticiens et étudiants.

# Thématique transversale

## Construction du sens

*Coordonnateur* : S. Osu

*Enseignants-chercheurs* : V. Abasq, A. Cambourian, G. Col, G. Deléchelle (e.r.), A. Delplanque (e.r.), J.-M. Fournier, T. Gallèpe, N. Garric, I. Girard, J. Longhi, D. Maurel, A. Piel, J.-P. Régis (e.r.), F. Toupin.

*Doctorants* : J. Constanza, P. Fournier, M. Martin, A. Ndione, E. Lecuit, J. M. Nagy, G. Potjans, S. Vanhoutte, L. Perbost, E. Sauquet.

*Autres participants* : C. Agafonov, F. Chen, S. Moukrim et des étudiants de Master 2.

*Un rapport détaillé, développant les parties théoriques et récapitulant les actions, est joint en annexe. La présentation ci-dessous en retient les aspects les plus significatifs.*

La spécificité, et l'originalité, de ce programme réside dans son objectif de faire collaborer à l'analyse de la construction du sens des spécialistes de disciplines linguistiques très différentes, soit non seulement des chercheurs habituels de l'énonciation (dans l'optique d'A. Culioli) mais aussi les chercheurs en TAL, en analyse du discours, en diachronie de l'anglais ou encore en morphophonologie, et ce dans des langues diverses dans un souci comparatiste double par nature : unité du langage, diversité des langues. En d'autres termes, on cherche ce qui varie de langue à langue et ce qui peut être considéré comme invariant, puis à quel niveau d'analyse (structurel ou abstrait) se situe un tel invariant. Il est ainsi né de la volonté des membres du LLL Tours de se retrouver autour d'un domaine d'action et d'interaction scientifique commun, à la fois lieu d'échange et de convergence avec l'ambition d'explorer les potentialités d'une linguistique plus intégrative.

### **Problématique : construction du sens et diversité des langues**

La plupart des recherches actuelles de la linguistique contemporaine intègrent d'une manière ou d'une autre, la dimension sémantique du langage en tentant d'une manière spécifique de rendre compte du sens en tant qu'entité produite dans et à travers le discours. Les travaux de la thématique « Construction du sens » s'inscrivent dans cette perspective en mettant l'accent non pas sur le résultat mais sur ce qui l'engendre. Suivant l'approche développée en linguistique de l'énonciation par A. Culioli (Culioli 1990, 1999 *et passim*), la construction du sens se trouve au centre des préoccupations de la linguistique qui entend décrire, caractériser et expliquer les phénomènes. L'objet « sens » est appréhendé comme le résultat de l'interaction de diverses opérations linguistiques représentées en surface par des formes appelées *marqueurs* (au sens de marqueurs d'opérations) actualisés par la prosodie, les mots, les syntagmes, les locutions ou les propositions dans des énoncés. Ces représentations en surface sont effectuées par le sujet énonciateur en fonction de la situation d'énonciation et toute recherche portant sur la construction du sens doit prendre en considération ces deux paramètres.

La recherche porte sur les langues étudiées par les chercheurs du LLL : allemand, anglais, dagara (Niger-Congo), espagnol, français, ikwère (Niger-Congo) et russe. L'approche est à la fois contrastive et typologique afin de cerner le mécanisme de la construction du sens à travers des phénomènes précis pour appréhender de manière spécifique comment chacune des langues représente des phénomènes particuliers en rapport avec la construction du sens.

## Réalisations et productions

L'équipe a abordé deux thèmes au cours des cinq années écoulées : la dénomination (2005-2007) et la temporalité (2008-2010).

### *Dénomination*

Le premier thème prolongeait les travaux précédents conduits sur les noms propres et sur les verbes d'apparence. La dénomination manifeste l'interaction entre le lexique (noms, noms propres), le discours (catégorie, catégorisation ou classification, construction d'identité), la morphologie (processus de formation des noms), la morphophonologie (e.g. la prononciation des noms propres dans les différentes langues), la syntaxe et l'énonciation (processus d'identification, négociation du sens). Le postulat de départ était que seule la prise en compte de l'interaction permettrait de rendre compte des phénomènes observés et de définir la dénomination. La première phase d'exploration s'est centrée sur le processus de catégorisation des êtres et/ou des objets dès lors que catégoriser, c'est identifier un élément à un groupe tout en marquant sa singularité, donc construire une identité à partir d'une entité (Delplanque, 2006). Les études sur la dénomination et sur la reduplication ont conduit à l'organisation de séances de travail (cf. annexes).

Concrètement, la dénomination a été abordée sous deux angles. Une première approche concerne le traitement automatique et le traitement phonologique des noms propres à partir de la variation de noms dans des textes traduits et du traitement phonologique des noms propres. Une seconde approche a pris en compte les conséquences de l'acte de re-nommer. La reduplication, distincte formellement et sémantiquement de la répétition, pose une forme (la base) et la reprend (la copie) soit partiellement soit entièrement ce qui revient à introduire ce que, dans le cadre de la TOPE, l'on appelle occurrence quelconque d'un domaine notionnel (base) et occurrence type d'un domaine notionnel (copie). La juxtaposition des deux formes se traduit en termes d'identification. Ainsi, l'occurrence quelconque est identifiable à l'occurrence type.

Ces travaux ont abouti à un colloque international : « Construction d'identité et processus d'identification » à l'Université François Rabelais (Tours, 29 & 30 novembre 2007) avec 66 communications et 2 conférences plénières (voir annexe). Les conférences plénières, filmées par le service audio-visuel de l'établissement, sont archivées sur le site de l'université :

[http://lettres.univ-tours.fr/1190190229151/0/fiche\\_\\_\\_actualite/&RH=1192087652523](http://lettres.univ-tours.fr/1190190229151/0/fiche___actualite/&RH=1192087652523)

39 contributions, sélectionnées par un comité scientifique ad hoc, ont donné lieu à la publication d'un livre de 623 pages : *Construction d'identité et processus d'identification* (P. Lang, Berne, 2010).

### *Temporalité*

En 2008-2010, le programme Temporalité a traité deux questions : (i) comment les langues et discours permettent-ils de construire le temps ? (ii) comment le facteur temps est-il utilisé dans la description linguistique ? A partir d'une approche contrastive, l'exploration des concepts de forme schématique, de repérage et de situation d'énonciation a servi de fil conducteur. En lien avec ce programme, a été organisée à Tours le 5 mars 2010, une journée d'étude sous la responsabilité de S. Osu et F. Toupin afin de mettre l'accent sur la perspective typologique des travaux et la pluralité des approches (cf. programme en annexe). L'ensemble des communications de la journée, enregistré par le service audio-visuel de l'Université François Rabelais, Tours, est archivé sur le site de l'université :

[http://lettres.univ-tours.fr/1266596192086/0/fiche\\_\\_\\_actualite/&RH=1179214274997](http://lettres.univ-tours.fr/1266596192086/0/fiche___actualite/&RH=1179214274997)

# Analyse de discours

## Composition de l'équipe

Coordonnateur : Nathalie Garric

Enseignants-chercheurs (EC associés) : J.-Y. Antoine, D. Maurel

Doctorants : J. Constanza

Collaborateurs : G. Bernier, F. Calas, V. Capdevielle-Mougnibas, L. Ghanam, M. Goldberg, A. Lefeuvre, H. Maurel-Indart, S. Point, M. Souchard, A. Tavernier.

Les activités de recherche de l'axe Analyse de discours (désormais AD) sont organisées en trois volets principaux complémentaires – un méthodologique, un épistémologique et un autre typologique – menés sous l'égide d'une conception de l'analyse de discours comme lieu de rencontre interdisciplinaire. Les travaux scellent cette orientation associant des chercheurs d'horizons variés, linguistique, littérature (F. Calas, H. Maurel-Indart), informatique (D. Maurel, J-Y Antoine), sciences de l'information et de la communication (M. Souchard, A. Tavernier), psychologie (V. Capdevielle-Mougnibas), sociologie et géographie (LISST-CIEU) ou encore biochimie (M. Goldberg) et sciences de la gestion (S. Point), mais tous confrontés aux textes et inscrits dans une problématique philologique. L'ambition de l'axe AD est, par ses collaborations interdisciplinaires plurielles, d'éviter la spécialisation des recherches dans une pratique sociale en particulier afin de développer une réelle réflexion méthodologique et théorique. L'objectif général est, plus que la connaissance d'une situation discursive, la théorisation de la discursivité. A cette fin, AD adopte pour objet d'étude des genres discursifs variés, certains comme le discours politique appartiennent à l'histoire de l'analyse de discours, d'autres résultent de nouvelles pratiques sociales en lien avec de nouveaux dispositifs communicationnels permis par les évolutions technologiques, d'autres encore s'apparentent à des genres transversaux tel que le discours propagandiste ou le discours scientifique médiatique par exemple.

## Méthodologie

Le volet méthodologique fait l'objet d'un investissement soutenu. Il consiste en une réflexion croisant les nécessités scientifiques de l'analyse de discours en lien avec l'extension et l'hétérogénéité des demandes d'intervention de son champ d'application avec les apports de la linguistique de corpus et des méthodes de traitement quantitatif des données textuelles. Les activités, dans ce cadre, se consacrent aux questions méthodologiques de recueil des données, de constitution des corpus et d'outillage de la textualité. Ces considérations ont reçu plusieurs formes de traitement.

L'une est menée en collaboration étroite avec V. Capdevielle-Mougnibas, membre du LPDPS (Psychologie du Développement et Processus de Socialisation, université de Toulouse le Mirail). Nous défendons l'intérêt de la mise en œuvre d'une méthodologie inductive basée sur un principe essentiel : la variation actualisable à tout niveau de la recherche. Nos travaux ont débuté dans le cadre de l'ACI (ACIEF/2004/1AO/20) « Education et Formation : contextes et effets » qui a été menée de Septembre 2004 à septembre 2007. Le principe de variation a ensuite été développé et systématisé par Capdevielle & Garric (2006, 2009), et par des séminaires mensuels organisés durant l'année 2010 par l'équipe LISST-CIEU sous la direction de I. Peyroux, auxquels V. Capdevielle et N. Garric ont régulièrement participé et ont été invitées à présenter leurs travaux.

Ce volet a également été mené par J. Longhi et N. Garric avec un intérêt tout particulier porté aux implications de la linguistique de corpus sur les objets et méthodes de l'analyse de discours. Le 24/01/2008, à l'université de Tours, l'axe AD a organisé sur cette thématique une journée d'étude : « L'analyse linguistique des corpus discursifs : des théories aux pratiques, des pratiques aux théories », dont les actes sont parus (Garric & Longhi, 2009). La question spécifique des données, de leur hétérogénéité et de la construction des corpus (écrites, orales, spontanées, suscitées, institutionnalisées)



a été spécifiquement abordée lors d'une conférence invitée de N. Garric à la MSH de Paris-Nord (Garric, 2010). La dimension applicative de l'analyse de discours a en outre été représentée lors des deux colloques consacrés aux applications et implications des sciences du langage (Léglise, Canut, Desmet & Garric, 2007). On soulignera, en relation avec cette problématique, le travail de (Mariscal, 2008).

Enfin, ce volet a été développé en collaboration par N. Garric et H. Maurel-Indart (EA2015) dans le cadre d'un projet PEPS-2009 « Analyse textuelle informatisée pour l'identification du plagiat : similitudes et différences, écarts et distance ». Celui-ci s'est notamment décliné sous la forme d'une journée de formation organisée à l'université de Tours à l'utilisation du logiciel ALCESTE, de deux journées d'étude les 10 et 11 décembre 2009 intitulées « Le style et sa formalisation », et d'une communication « Autour de la formalisation du style ». La publication des actes de ces journées est actuellement en cours, elle sera finalisée au cours des prochains mois.

Le volet méthodologique de l'axe analyse de discours trouve enfin des exploitations pédagogiques. Des séminaires abordant directement ses préoccupations, « Méthodologie de la recherche et de l'analyse en situation » et « Outils informatiques pour l'analyse linguistique », ce dernier associant des enseignants du LLL et du LI, sont dispensés dans le cadre du master. Au sein de ces séminaires, nous maintenons une politique de contact entrepreneurial afin de confronter directement les étudiants au terrain de l'analyse de discours. Plusieurs étudiants ont ainsi pu effectuer des stages dans des entreprises (EDF, INIT, Tendances Institut).

## Epistémologie

La réflexion épistémologique, indirectement abordée par le volet méthodologique, est développée de façon originale par une analyse du discours de la médiation télévisuelle. Nous postulons, en effet, que le dispositif qu'inaugure *L'Hebdo du médiateur* constitue, de par le principe d'institutionnalisation de la parole citoyenne participative qui le sous-tend et probablement également son fonctionnement discursif stratégique, une voie d'accès au genre discursif médiatique, en particulier au sous-genre d'information télévisuelle, et plus généralement à la discursivité elle-même. L'émission, en tant qu'espace d'observation pour l'analyste du discours mais aussi de réflexion ou d'auto-réflexion pour ses acteurs, est conçue comme lieu de divulgation de la discursivité susceptible d'éclairer certains fonctionnements langagiers en les montrant. Les pratiques discursives, du fait de leurs spécificités, du fait de la posture épидiscursive qu'elles suggèrent, sont posées comme des « participants épistémologiques » à l'analyse du discours elle-même, mis au service de la définition de ses objets, de ses catégories descriptives ou encore de ses concepts théoriques. Cette réflexion a donné lieu à plusieurs travaux : (Garric, à par., 2009) et à des conférences à Helsinki (2007) et à l'Université de Laval (2007).

Le volet épistémologique a également été développé, en lien avec la problématique transversale sur la construction identitaire du LLL, par un intérêt spécifique pour les opérations de nomination, de dénomination et de catégorisation. La participation de J. Longhi, N. Garric mais également de plusieurs doctorants, J. Constanza, ou étudiants de master, L. Ghanam, A. Lefeuvre, G. Bernier, inscrits en analyse de discours, au colloque « Construction d'identité et processus d'identification » organisé à l'université de Tours par le LLL témoigne de l'investissement de cet axe dans cette question relative à la construction du sens discursif. Les actes du colloque (Peter Lang, 2010) intègrent plusieurs de ces contributions. La nomination, et partant la notion d'identité, sont questionnées par deux étudiants inscrits en thèse. E. Lecuit, qui devait être recrutée en CIFRE par Synapse développement, projet abandonné après un changement de direction de l'entreprise, s'est inscrite en doctorat pour étudier en linguistique contrastive les entités nommées (Lecuit, Maurel, Vitas & Krstev, 2009). J. Constanza, sous la direction de F. Nemo et N. Garric, étudie en thèse l'emploi des noms propres dans le discours politique (Constanza, 2010).

La dimension épistémologique est également développée par les travaux menés en linguistique textuelle dans le cadre du projet PEPS et de sa poursuite sous la forme du projet DIATEXT et par la recherche

de M. Nagy, doctorant en cotutelle sous la direction de T. Gallèpe (LLL-Tours) et de J-M. Adam (Université de Lausanne), sur la linguistique textuelle d'E. Coseriu. La qualité des travaux menés par deux étudiants inscrits dans l'axe AD a permis leur sélection sur des projets de thèse financés. A. Lefeuvre a été accueillie au LaBRI de Bordeaux pour travailler sur « l'extraction d'itinéraire dans les récits de voyages » et V. Mariscal est en contrat doctoral dans le laboratoire Valibel sous la direction de P. Hambye de l'UCL.

## Typologie

Le volet typologique relève d'une problématique générique descriptive qui s'appuie sur l'analyse empirique de corpus variés pour identifier les caractéristiques discursives propres à différentes situations de communication sociales. L'axe AD développe dans ce cadre des études du discours politique menées par J. Constanza et V. Mariscal, du discours écologique dont N. Garric et V. Mariscal ont rendu compte en 2009, « L'écologie politique dans les livres-programme électoraux » et du discours patronal et/ou économique (Léglise & Garric, 2006, 2008). Le volet typologique se consacre également à l'analyse du discours scientifique (Garric & Goldberg, à par.).

Enfin, le laboratoire LLL s'est consacré également au discours expert en s'associant aux séminaires du GTAD (MSH, Paris) sur cette thématique. N. Garric et I. Léglise ont organisé durant l'année 2008 deux journées d'étude et travaillent actuellement à une publication consacrée à ce thème de recherche. Les contributions à l'ouvrage se situent dans le champ de l'analyse de discours, avec une attention particulière à l'étude de formes linguistiques apparaissant en corpus, mais également dans divers champs disciplinaires des sciences humaines et sociales s'intéressant également aux discours d'expert. Dans le champ de l'analyse du discours, les textes se fondent sur des analyses de corpus mettant en jeu la place de l'expertise ou la figure de l'expert, légitimé institutionnellement ou non. Une triple attention sera donc marquée quant à la spécificité du champ indexical que le corpus représente (médias, pénalité, débats publics, etc.), quant aux outils analytiques mobilisés (analyses quantitatives, lexicométriques, argumentatives, conversationnelles etc.) et quant aux enjeux sociaux, politiques ou idéologiques que l'analyse permet de caractériser ou d'informer. Les contributions d'autres disciplines abordent le thème de l'expertise et la position de l'expert selon un rapport aux sources expertes qui n'emprunte pas les voies de l'analyse linguistico-discursive. L'ambition de ce projet est de faire communiquer les apports des sciences humaines et sociales avec les analyses de corpus attentives aux contraintes et effets sociaux dont le langage est porteur. Suivant les champs du savoir concernés par l'analyse (sociologie, droit, philosophie, histoire, sciences de l'information et de la communication, anthropologie etc.) les textes abordent la spécificité de la place sociale de l'expert ou de l'expertise : constitution de la figure d'expert, variété, croisements, antagonismes entre différentes formes et différents acteurs de légitimité et de pouvoir, etc.

# Morphophonologie

## Composition de l'équipe

Coordonnateur : J.-M. Fournier

Enseignants-chercheurs : V. Abasq

Autres participants : N. Ballier, A. Delplanque, J.-L. Duchet, I. Girard, S. Hanote, S. Moore, S. Osu, N. Rossi-Gensane, I. Trevian, F. Zumstein.

Doctorants : E. Descloux, P. Fournier, M. Hassen, M. Martin, S. Vanhoutte

Ce programme s'est donné pour objectif, à partir du traitement des bases dictionnaires, d'étudier l'interface de la phonologie et de la morphologie en anglais, en particulier dans la relation entre les notations orthographiques et les interprétations prosodiques.

En 1979, Lionel Guierre initiait un paradigme d'analyse de la prononciation de l'anglais présentant un certain nombre de particularités dont certaines sont restées originales dans l'analyse phonologique contemporaine. En premier lieu il étayait ses travaux sur un traitement informatique systématique d'une base de données lexicales (informatisation de la 12<sup>e</sup> édition du *Pronouncing Dictionary* de Daniel Jones). En second lieu ses travaux mettaient l'accent sur l'interface entre phénomènes phonologiques et morphologie d'abord, mais aussi syntaxe (catégories) et sémantique, une perspective initiée par Chomsky et Halle, mais guère poursuivie depuis par les diverses écoles anglo-saxonnes. Enfin, ils présentaient la particularité, controversée, d'intégrer la prise en compte de la forme écrite.

C'est dans ce paradigme que se sont inscrits dès l'origine les travaux conduits à Tours, en y intégrant l'articulation diachronie/synchronie à laquelle l'anglais se prête de façon toute particulière en raison de son parcours spécifique au sein des langues indo-européennes, à l'articulation des branches romane et germanique. Ce versant de la recherche est conduit en collaboration avec le FORELL (Université de Poitiers). Le recours à des corpus et un traitement informatiques s'est trouvé encouragé par une collaboration suivie avec le LI (Laboratoire d'Informatique de Tours) ; la constitution d'une équipe de recherche en linguistique à l'Université de Tours a en outre conduit à un élargissement des recherches vers d'autres langues, notamment le français en collaboration avec N. Rossi.

Plus particulièrement centrées sur la problématique de l'accent, lexical en anglais, les recherches accordent une place privilégiée à l'interaction des niveaux linguistiques, en particulier entre phonologie et morphologie/lexique, ainsi qu'en témoignent les thèses engagées ou soutenues dans ce domaine :

Abasq, V. : *Préfixation et particules adverbiales en anglais contemporain : étude du comportement accentuel* (soutenue en décembre 2007, aujourd'hui MCF à l'Université de Tours)

Girard, I. : *Isomorphisme en anglais contemporain. Etude de quelques suffixes* (soutenue en décembre 2007, aujourd'hui MCF à l'Université de Boulogne sur mer)

Descloux, E. : *Prononciation de la voyelle accentuée des mots en 'C<sub>0</sub>V<sub>0</sub>CV<sub>0</sub>C en anglais contemporain* (thèse financée par bourse régionale de recherche)

Fournier, P. : *Prononciation des suffixés en -ous en anglais contemporain*

Hassen, M. : *Les composés savants en anglais contemporain*

Martin, M. : *Phonétique et phonologie de l'anglais australien contemporain*

Vanhoutte, S. : *Étude linguistique des digraphes en anglais contemporain*

S'attachant à l'analyse de la structure des phénomènes et à leur modélisation, les recherches opèrent un recours systématique aux outils informatiques, ce qui conduit à fonder la réflexion et la production sous quatre aspects complémentaires :

- Bases de données
- Simulations
- Modélisation et tests des hypothèses
- Outils d'analyse et didacticiels

C'est dans ce cadre que le *LLL* a initié en 2009 le projet de constitution d'une base de données dictionnaire sur la prononciation de l'anglais contemporain, projet auquel s'est associée l'Université de Poitiers, à partir de l'exploitation de bases de données informatiques, notamment la 12<sup>e</sup> édition du dictionnaire de Daniel Jones, *English Pronouncing Dictionary*, réalisée par Lionel Guierre dans les années 60, et la première édition du dictionnaire de J. C. Wells, *Longman Pronunciation Dictionary*, de 1990, mis à l'époque à la disposition des chercheurs français par son auteur.

La relative ancienneté des bases de données (le dictionnaire de Jones en est à sa 17<sup>e</sup> édition et celui de Wells à sa 3<sup>e</sup>) implique de la part des chercheurs un travail lourd et fastidieux de confrontation manuelle avec les dictionnaires plus récents tandis que le format et les informations contenues dans ces bases de données ne permettent pas de systématiser les analyses de façon satisfaisante. La constitution d'une base de données à jour réalisée à partir de ces deux dictionnaires, désormais accessibles au format xml, sera enrichie avec les informations morphologiques, lexicales, syntaxiques et sémantiques indispensables à l'analyse des comportements phonologiques, mais absentes de ces dictionnaires. L'engagement d'une recherche doctorale sur l'anglais australien, peu étudié de ce point de vue, a conduit à un partenariat avec les éditeurs du *Macquarie Dictionary (of Australian English)* qui assure à l'équipe la mise à disposition de leur base en contrepartie de l'expansion de leurs variantes de prononciation.

Contrairement à la partie historique du projet général réalisée à Poitiers, les accords de licence avec les éditeurs interdisent de rendre publique la base de données qui sera réalisée. En revanche, les résultats de la recherche et la valorisation restent acquis au laboratoire. Les partenariats négociés avec les éditeurs des trois dictionnaires concernés paraissent prometteurs et laissent augurer des perspectives intéressantes pour les étudiants impliqués dans ce programme qui comprend :

- constitution d'une base de données relationnelle à partir des trois dictionnaires contemporains,
- normalisation de la base et de la base poitevine pour assurer l'interfaçage.
- élaboration de programmes de traitement et d'analyse automatiques.

Comme il n'existe pas à ce jour de base de données sur la prononciation de l'anglais contemporain à la mesure de celle en cours, ce travail doit permettre de développer, tester et mettre en œuvre l'analyse du fonctionnement de la prononciation de l'anglais avec une fiabilité inégalée sur la scène scientifique internationale. L'ampleur de la tâche implique d'inscrire ce projet dans un programme général de longue durée, dont les deux premières années (2010-2011) seront consacrées à la réalisation de la base et à un début d'exploitation. L'interfaçage avec la base de données historique constituée à Poitiers permettra de vérifier et consolider cette analyse en l'articulant à la dynamique diachronique : ici encore, si cette voie a été parfois explorée, elle ne l'a jamais été avec le caractère systématique et empiriquement vérifié qui caractérise notre projet scientifique.

Les deux domaines sont pris en charge par :

- le FORELL : dans le prolongement des recherches déjà conduites sur la variation en anglais contemporain et sur les éditions électroniques déjà réalisées des dictionnaires de Bailey (1727), Buchanan (1766) ou partiellement réalisée (Stephen Jones, début XIX<sup>e</sup> s.), afin d'établir une édition numérique du dictionnaire de John Walker, *A Critical Pronouncing Dictionary And Expositor of the English Language* (2<sup>e</sup> éd, 1797), qui cumule les observations de nombreux autres auteurs du XVIII<sup>e</sup> siècle.

- le *LLL* : réalisation de la partie corpus dictionnaire contemporain à partir des bases de données des deux dictionnaires de référence sur la prononciation britannique contemporaine (incluant des données sur la prononciation américaine) : *Longman Pronunciation Dictionary* et *Cambridge English Pronouncing Dictionary*, et celle du dictionnaire de référence de l'anglais australien, *Macquarie Dictionary*. Elaboré en base de données relationnelle, le corpus complètera les informations de ces dictionnaires, en particulier les deux premiers, dans les domaines de la morphologie, la lexicologie, la syntaxe et la sémantique.

Le calendrier prévisionnel est établi de la façon suivante :

- finalisation des négociations avec les éditeurs (septembre 2009)
- acquisition numérique du dictionnaire de Walker en images de page (automne 2009)
- conception et réalisation du logiciel de gestion des bases (2010-2011)
- formatage des dictionnaires déjà numérisés (juin 2010)
- reconnaissance optique et post-édition du dictionnaire de Walker (décembre 2010)
- balisage des données et constitution d'une base de données panchronique (2011)
- exploitation et développement (2012-2015)

Il est à noter que ce projet a fait l'objet d'un financement des Universités de Tours et de Poitiers pour sa phase initiale.

Au nombre des activités liées à ce projet, on compte une dizaine de conférences en colloque (Aix-en-Provence, Brisbane (Aus), Chambéry, Edinburgh, Montpellier, Orléans, Paris, Aix), des interventions lors de trois journées d'études, des articles et le *Manuel d'anglais oral* de J.-M. Fournier (Ophrys, 2010).

Ce projet s'intègre dans la relance, par les universités de Tours et d'Orléans, du Réseau Français de Phonologie où une communication sur ce thème a été présentée.

# LTI

## Langue et Traitement de l'Information

Coordonnateur : D. Maurel (associé)

Enseignants-chercheurs : J.-Y. Antoine (associé), I. Eshkol, N. Friburger (associée), N. Garric

Doctorants : L. Ghanam, E. Lecuit

Collaborateurs :

### Historique

Depuis la création du groupe de recherche Langues et Représentation, il existait un axe intitulé *Langue et traitement de l'information (LTI)*, composé d'enseignants-chercheurs intéressés par le recours à l'informatique, soit dans le cadre de l'analyse du discours, soit dans celui du traitement automatique des langues. A cet axe étaient associés (entre autres) les collègues de 27<sup>e</sup> section, en rattachement principal au laboratoire d'informatique (LI) de l'université de Tours. Plusieurs travaux et publications en commun avaient été réalisés, en particulier la participation au projet Technolangue « NomsPropres ». La mutation de deux EC du LLL, N. Rossi et Th. Grass et le réinvestissement de N. Garric dans l'axe AD a eu pour conséquence de ne laisser dans LTI que les collègues associés qui ont engagé des collaborations avec l'axe AD et la dynamique entre Orléans et Tours, suite à la création du LLL, a suscité de nouvelles collaborations avec Orléans (opérations ANR Variling et CO2). Si l'axe LTI ne sera pas reconduit, les collaborations engagées se poursuivront.

### Collaborations avec l'axe *Analyse de discours*

Dans le cadre de l'axe de recherche transversal sur la notion d'identité, et dans la continuité des travaux antérieurs de LTI, deux sujets de thèse, co-encadrés par D. Maurel et N. Garric, devaient être définis avec des entreprises, fin 2008, dans le cadre de bourses CIFRE. Elles sont présentées dans la partie AD.

### Collaboration avec l'opération Variling

La coopération entre chercheurs du LLL et chercheurs associés du LI s'est concrétisée par la participation au programme ANR Variling. Le travail réalisé dans ce projet correspondait à la recherche et au balisage des entités nommées et dénommantes sur le corpus ESLO, ainsi que sur l'étude des disfluences aux alentours de ces entités. Ce travail a été en grande partie réalisé à l'aide du logiciel CasSys, développé par Nathalie Friburger, qui utilise le principe d'une description linguistique par une cascade de transducteurs Unitex. Une cascade de reconnaissance des entités nommées, réalisée pour du texte journalistique, a été adaptée au corpus oral transcrit ESLO avec un balisage XML inspiré de la campagne d'évaluation ESTER. Puis a été créée une cascade des « entités dénommantes », c'est-à-dire des entités permettant une description fine du locuteur ou de sa famille avec un double objectif : mettre le balisage à disposition d'études sociologiques et aider à l'anonymisation du corpus. Les cent douze fichiers ont été relus (et corrigés manuellement) à partir des balisages effectués. Cette ANR a été abondée à Tours par un complément financier du FEDER Région Centre qui a permis la participation à la campagne d'évaluation ESTER et le balisage d'une partie importante de la base Frantext (4 767 fichiers), à la demande du CNRTL-Nancy.

Ces études ont conduit à une série de publications : Eshkol, Gazeau & Maurel, 2008 ; Maurel, Friburger & Eshkol, 2009 ; Eshkol, Maurel & Friburger, 2010 qui témoignent d'une véritable activité conjointe dans ce domaine.

## **Collaboration avec l'opération CO2**

Les universités de Tours et d'Orléans ont lancé conjointement en 2009 un appel d'offre pour financer des projets communs. Le LLL et le LI ont été financés pour deux ans sur un projet concernant l'étude des coréférences dans le corpus ESLO, le projet CO2. Le projet a débuté par une étude comparée des logiciels existants permettant l'annotation XML des coréférences. Il se poursuit par un balisage des groupes nominaux, qui a été entrepris à l'aide de CasSys, afin de faciliter la future tâche d'annotation des coréférences. Le projet donnera lieu à diffusion de ressources. Il est conçu comme une première étape vers une structuration des recherches francophones sur le sujet, sous le pilotage du LLL.

# Diachronie de l'anglais

## Composition de l'équipe

Coordonnateur : F. Toupin

Enseignants-chercheurs : J.-P. Régis (jusqu'en 2009), S. Gatelais

L'équipe « diachronie de l'anglais » représente dans le LLL le versant historique de la linguistique de l'anglais. Ce domaine, très peu structuré en France, a connu un développement récent dont témoignent la mise en place d'un colloque périodique et des publications.

## Le Colloque Bisannuel de Diachronie de l'Anglais

Le Colloque Bisannuel de Diachronie de l'Anglais (CBDA) est co-organisé par le LLL à Tours et par le laboratoire Corpus-Amiens (EA 4295). Sa première session a eu lieu les 5 et 6 juin 2008 avec deux conférences plénières d'A. Rousseau et J.-M. Gachelin. Une journée a été consacrée à des ateliers de travail par niveau (analyse linguistique (marqueurs en contexte, questions d'ordre des mots, etc.), traduction, paléographie) sur des textes en vieil- et moyen-anglais avec, en conclusion, la cérémonie de remise des mélanges à André Crépin à l'occasion de son 80<sup>e</sup> anniversaire. La seconde journée a été consacrée aux communications dont plusieurs présentées par de jeunes collègues. Une assistance choisie, à la dimension d'une spécialité faiblement représentée dans les départements d'anglais, a conduit les organisateurs (Toupin/Lowrey/Delesse) à renouveler et étendre l'expérience en préparant le prochain colloque à Tours en 2011.

La publication des actes a contribué à donner une nouvelle visibilité à ce domaine en France (Delesse, Lowrey & Toupin, 2010), où se trouve confirmée la place de l'Université de Tours dans un champ où le LLL a été représenté par une série d'articles (Toupin, 2008, 2009, 2010 publications, Toupin & Lowrey 2009), de communications (Toupin 2008, 2009) et par l'animation d'un atelier à partir de la traduction commentée de 2 textes alfrédiens : *Bald's Leechbook* et *Alfred's Preface to the Pastoral Care* lors du 1<sup>er</sup> CBDA et par l'organisation de l'atelier « Chaucer's language: a creole? » lors du 15<sup>e</sup> colloque de la New Chaucer Society à New York en 2006. Cette activité a été complétée par la relecture d'articles pour l'ALAES, CORELA et la revue *English Language and Linguistics*.

La reconnaissance du travail accompli en ce domaine et son nécessaire développement ont conduit le CED 11<sup>e</sup> section de Tours à donner sa préférence au recrutement d'un linguiste ayant, dans ses compétences, une bonne connaissance de la diachronie de l'anglais, S. Gatelais.



## CORELA

CORELA (Cognition, Représentation, Langage) est la revue du Cercle Linguistique du Centre et de l'Ouest. C'est une revue semestrielle et en accès libre ; CORELA est hébergée à l'Université de Poitiers.

L'originalité de CORELA est son caractère non thématique. Notre choix éditorial est de publier des contributions sur des thèmes ouverts, qui abordent des questions de langues (grammaire, lexique, syntaxe, phonologie, etc.), mais aussi de langage, de représentation, de modélisation, de sémantique. Nous souhaitons accueillir des contributions abordant le plus grand nombre des domaines de la linguistique. Nous essayons, dans la mesure du possible, de regrouper des articles proches par leur approche, leur méthodologie, leur corpus, ... sans toutefois chercher à imposer une thématique.

CORELA permet également la diffusion de travaux de jeunes chercheurs, ou de chercheurs confirmés sur des thématiques "marginales" ou complémentaires de la série des Travaux Linguistiques du CerLiCO (actes du colloque annuel de l'association CerLiCO, publiés aux Presses Universitaires de Rennes).

Le choix du support électronique adopté par CORELA répond à une volonté de confronter plus facilement et plus rapidement des points de vue ainsi que d'augmenter la périodicité des publications du CerLiCO.

CORELA comporte également une base de données qui regroupe des contributions sous un autre format que des articles, dans une collection de numéros spéciaux (actes de colloque ou de journée d'études, compte rendu de manifestations scientifiques ou d'articles, pré-publications, textes de conférences, annonces de colloques, activités doctorales, ...). Nous ne publions pas de thèses en l'état ; les auteurs peuvent cependant soumettre un article basé sur leurs recherches doctorales.

L'objectif de CORELA est d'encourager des discussions scientifiques et théoriques les plus larges possibles. Le choix du support électronique est d'en faciliter la publication et la diffusion. Il est aussi de faciliter l'accès aux publications déjà en ligne et de faire d'une revue en ligne une banque de connaissances aussi bien qu'un lieu de discussion.

CORELA est dirigée par G. Col, assisté de J.-M. Fournier avec, au comité scientifique, de G. Deléchelle et D. Maurel et au comité de lecture A. Delplanque, N. Garric et S. Osu.

Consulter : <http://corela.edel.univ-poitiers.fr/>

**Pour les publications 2007-2010, voir annexes**

## Revue de Sémantique et Pragmatique (RSP)

Chaque domaine de la linguistique tend progressivement à élaborer son propre programme de recherches et définit par cette spécialisation son identité et les moyens de sa reconnaissance.

La détermination d'une compétence nouvelle – celle imposée depuis une trentaine d'années par la pragmatique linguistique – ou réactualisée – les nouvelles sémantiques – s'accompagne de débats. Ceux-ci portent sur les concepts, la formalisation, la constitution de l'objet et la validation de méthodes, définissant des écoles dont les débats et les oppositions terminologiques remettent à chaque instant en cause l'unité précaire du champ.

Depuis que les disciplines issues de la philosophie du langage ordinaire et du formalisme logique ont établi leurs premiers résultats, la diversité des propositions et des équipes a rendu difficile la continuation des échanges et une vision d'ensemble. Aucune revue francophone ne proposant une analyse de ces thèmes, les auteurs n'ont jusqu'à présent le choix qu'entre la dispersion dans des revues généralistes ou des publications en langue étrangère. Or, dans ces deux domaines, les traditions académiques comptent beaucoup et l'originalité de l'école française n'a pu s'affirmer à hauteur de ses travaux faute d'un support qui en refléterait une image auprès de l'ensemble de la communauté.

La *Revue de Sémantique et Pragmatique* entend ménager cet espace ouvert à la diffusion de la recherche, sans exclusive.

Consulter : <http://www.univ-orleans.fr/RSP/>

**Pour les publications 2007-2010, voir annexes**

## CAT COD

A l'initiative d'E. Schang et avec l'aide de

Jean-Yves Antoine	(LI (associé LLL) Université Tours)
Claude Barras	(LIMSI-CNRS Paris XI)
Steven Bird	(University of Melbourne & LDC University of Pennsylvania)
Lou Burnard	(Oxford University Computing Services)
Pascal Cordereix	(BNF, Paris)
Benoît Habert	(ENS-LSH, Lyon)
Serge Heiden	(ENS-LSH, Lyon)
Nancy Ide	(Vassar College)
Michel Jacobson	(Ministère de la Culture et de la Communication, Paris)
Laurent Romary	(MPI Berlin-INRIA)
Emmanuel Schang	(LLL, Université d'Orléans)
Richard Walter	(CNRS, Orléans)
Peter Wittenburg	(Max-Planck-Institute for Psycholinguistics, Nijmegen)

a été organisé le premier colloque CAT COD à Orléans en 2008.

### *Argumentaire*

Jusqu'à récemment, le nombre d'enregistrements oraux numérisés et disponibles pour l'étude et la description des langues était relativement faible et ces derniers étaient confinés dans des agences spécialisées qui en assuraient le partage. Avec l'essor du web et des capacités de stockage, de diffusion et de traitement, il est devenu abordable pour des structures telles que les laboratoires de recherche qui peuvent diffuser elles-mêmes leurs ressources orales. La capitalisation des efforts de numérisation des données linguistiques est devenu possible afin d'établir les bases empiriques de nouveaux programmes de recherche.

L'observation des normes de codage et de catalogage de ce type de ressources dans les différents projets existant en France et en Europe, montre une grande hétérogénéité des pratiques. Alors que ces mêmes projets se sont montés dans le but de faciliter l'accès, le partage ou la conservation des données, on constate que cette diversité des formats, des codages et des protocoles utilisés limitent justement ces objectifs.

### *Catalogage*

La croissance récente et rapide du nombre d'enregistrements oraux disponibles sur le web demande à être accompagnée d'un effort important de description et de référencement afin que ces données soient accessibles facilement, ne soient pas noyées dans la masse et que la gestion en soit facilitée. Il convient à ce titre de distinguer des pratiques de catalogage qui ont pour vocation le suivi du cycle de vie d'un projet de constitution de ressources, de celles qui ont pour vocation à guider l'exploitation, la conservation et la diffusion de ces ressources.

Cette activité de catalogage est d'autant plus importante et urgente que les opérations de manipulation sur cette masse de données augmentent elles aussi (échange, maintenance, recherche, etc.). Certaines communautés se sont déjà fortement organisées et ont parfois établi des normes comme le Dublin Core pour ce qui concerne les ressources sur le web, ou depuis plus longtemps les normes MARC maintenues par la Bibliothèque du Congrès pour la description des ressources bibliographiques.

Plus récemment, des communautés plus restreintes en linguistique ont établi des propositions de catalogage (OLAC, IMDI). Il existe maintenant suffisamment d'expériences dans l'utilisation de ces propositions pour en faire la critique, proposer des améliorations, des pistes de réflexion et pour établir une charte minimale à respecter par ceux qui souhaitent s'engager dans la diffusion de ressources orales linguistiques afin d'en faciliter l'échange et plus généralement l'utilisation scientifique.

### *Codage*

Alors que le catalogage est essentiel à l'identification des ressources et à leur comparaison, le codage est central dans la description (donc l'interprétation) du contenu de la ressource à des fins d'exploitation pour des études spécifiques.

Si le codage d'une transcription d'un corpus sonore équivaut à l'explicitation d'une interprétation établie dans la perspective propre d'une recherche, l'extrême diversité des pratiques a été mise en évidence par les recensions effectuées lors du lancement des programmes EAGLES, MATE puis ISLE qui ont montré à quelles difficultés se trouvait confrontée la communauté active en documentation automatique et linguistique. La spécification d'un outil logiciel d'annotation universel suppose l'établissement d'un consensus pour la catégorisation, la dénomination et la structuration des phénomènes afin de parvenir à une véritable interopérabilité des données en vue d'exploitations multiples et qui n'obèrent pas des investigations futures encore peu prévisibles quant aux formes qu'elles prendront mais évidentes dans leur constitution. Ce travail de standardisation dans le codage des corpus oraux a déjà été engagé dans le domaine des corpus textuels avec la TEI.

Cette étape réalisée, quand un compromis sur le catalogage et le codage des données orales aura été négocié, il restera à fixer les règles et à construire les outils de vérification concernant la conformité des données particulières avec les principes établis. Tels sont les objectifs que s'est assignés l'expertise de CAT-COD en France au cours de son premier colloque et d'initiatives à suivre sous forme de journées d'études et de publication électronique.

L'association avec la BnF devrait donner à ce programme un nouvel élan.

## Réseau Français de Phonologie

Prenant la suite du GDR « Phonologies », le Réseau Français de Phonologie (RFP) a contribué pendant une dizaine d'années, par l'organisation de colloques, de journées d'études et de séminaires, au dynamisme de la discipline. Après un dernier colloque à Aix en 2005 (le précédent s'était tenu à Orléans en 2004), les activités se sont interrompues.

Des collègues du LLL (G. Bergounioux, J.-M. Fournier), avec l'aveu de l'ensemble des participants, en particulier de B. Laks qui avait créé et animé cette structure, ont pris l'initiative d'une relance afin de préparer un nouveau cycle de rencontres. Ils ont réuni l'ancien comité scientifique et élargi à de jeunes collègues. (voir annexe)

Comme l'indiquait l'appel à communications :

*Toutes les écoles et tous les courants de la phonologie sont invités à participer. Les problèmes traités peuvent concerner la phonologie générale comme l'analyse phonologique d'un phénomène linguistique ou d'une langue particulière. Seront prises en considération les soumissions concernant l'acquisition, la diachronie, la dialectologie, les formalismes, l'extraction du signal, l'épistémologie et toutes les questions qui s'inscrivent explicitement dans le domaine de la phonologie et de ses interfaces.*

Il s'agit d'un colloque sans publication mais le RFP s'appuie sur une liste de diffusion (centralisée à Orléans), un site (géré à Lille 3 par Roland Noske). La structure d'organisation est très souple : les laboratoires financent les inscriptions et les voyages de leurs chercheurs et l'université d'accueil prend en charge les dépenses courantes. La coordination du RFP est assurée pour un an par l'organisateur du colloque, G. Bergounioux en 2009-2010 pour le colloque qui s'est tenu à Orléans, J.-M. Fournier en 2010-2011 pour Tours, J. Brandao de Carvalho et S. Wauquier ensuite pour Paris.

Pour le programme du RFP, voir annexes

## VARILING

Variling est un programme coordonné par le LLL qui s'est associé pour le conduire à bien avec le Laboratoire d'Informatique de Tours (D. Maurel), le Centre de Linguistique Théorique de l'EHESS (M. de Fornel) et le laboratoire Modyco (Paris X).

Anticipant la démarche entreprise à l'occasion de ce contrat, le LLL figurait dans cet appel Corpus 2006 avec l'intention de tester en grandeur réelle une démarche de maîtrise intégrale du processus de constitution d'une base de données sonores, du protocole d'enquête à la mise à disposition sur Internet.

Comme l'indique le titre, la question de la variation était centrale dans l'approche, justifiant une comparaison diachronique (ESLO1 et ESLO2) mais aussi géographique avec la prise en compte de travaux effectués à Bordeaux par l'équipe du CELITH. Le traitement spécifique des données, à la fois dans le traitement des disfluences et dans l'anonymisation, justifiait la collaboration avec Tours.

Le programme arrive à échéance en janvier 2010 (une prolongation d'un an a été sollicitée). Les rapports remis à l'ANR feront le point sur les réalisations qui sont pour une grande part accessibles à tous grâce au contrat conclu avec la société ARES pour la mise en ligne des données.

On trouvera dans les annexes les grandes lignes du projet tel qu'il a été conçu et, ce qui est plus rare, en grande partie réalisé. Dans la continuité du travail collectif effectué, et en profitant de l'apport de l'expertise de la BnF, cette recherche doit se poursuivre au cours du prochain quadriennal, en partenariat avec les autres projets Corpus et sous l'égide de l'ILF pour aboutir à la constitution d'un grand corpus du français parlé.

La présentation scientifique du projet est reproduite dans les annexes.

*Concernant la mise en ligne de la plate-forme réalisée par ARES, elle était en cours de validation par le service informatique de l'Université à la date de rédaction de ce document. Une adresse http sera communiquée aux rapporteurs.*