



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^{er} 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^e et 11^e 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^e section à Orléans, retraite, 1 en 11^e section à Tours, retraite, et 2 en 12^e section à Tours, mutation) ;

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

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

- 1 PR en 11^e section à Tours,

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

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

Un onzième EC (70^e/7^e) 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, wayampì) 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^e 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^e PROHEMIO (El Rebollar 2008 et 2009),
5^e 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, wayãpi : 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^e-XV^e 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^e 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^e *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^{er} 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 inexistência 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^e) dans une perspective curriculaire », XI^e 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 logiciel 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^e/3^e) 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^e année

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

2^e 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 PROHEMIOS'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^e 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^e 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^e. 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, ikwere (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

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

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^e 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 'C₀V₀CV₀C 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^e 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^e édition et celui de Wells à sa 3^e) 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^e s.), afin d'établir une édition numérique du dictionnaire de John Walker, *A Critical Pronouncing Dictionary And Expositor of the English Language* (2^e éd, 1797), qui cumule les observations de nombreux autres auteurs du XVIII^e 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^e 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^e 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^{er} CBDA et par l'organisation de l'atelier « Chaucer's language: a creole? » lors du 15^e 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^e 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 (associié 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.

LABORATOIRE LIGÉRIEN DE LINGUISTIQUE

Université d'Orléans / Université de Tours
Institut de Linguistique Française (CNRS)
EA 3850 > UMR ?

Ecole doctorale « Sciences de la société : norme, régulation, organisation, territoire »
(Orléans – Tours)



Au cours du prochain contrat, le LLL ambitionne sa transformation en une UMR associant :

- au Laboratoire Ligérien de Linguistique (Université d'Orléans & Université de Tours),
- le Département de l'Audiovisuel de la Bibliothèque nationale de France
- le CNRS / INSHS section 34 : « Langues, langage, discours ».

Les négociations avec la BnF et le Ministère de la Culture et de la Communication n'étant pas finalisées à la date de rédaction (septembre 2010), le dossier a été conçu comme un cadre de référence où se trouvent exposés l'argumentaire scientifique et, dans ses grandes lignes, le fonctionnement administratif envisagé. A définition thématique et ancrage institutionnel constants, les modalités organisationnelles pourront être précisées dans leur formulation, en fonction des attentes des partenaires, pour devenir effectives à l'échéance contractuelle, soit le 1er janvier 2012.

Tel qu'il est, ce programme permet, sans attendre l'issue des discussions en cours, une évaluation par les instances concernées (AERES, BnF, CNRS, universités de rattachement, MESR et MCC) et la poursuite du dialogue avec les instances de collaboration, l'Institut de Linguistique Française (FR 2393), la Délégation Générale à la Langue Française et aux Langues de France (DGLFLF /MCC), l'Agence Universitaire de la Francophonie (AUF), le Centre de Ressources pour la Description de l'Oral (CRDO – Université de Provence), le TGE-ADONIS et les UMR de la section 34 (voire 35) que nous entendons associer à notre démarche sur les corpus oraux.

TABLE DES MATIERES

Résumé / Summary

1. Présentation de l'UMR LLL / BnF / CNRS

- l'aboutissement d'une démarche
- Universités : une politique de la recherche au niveau du PRES, et au-delà
- BnF : une valorisation des savoirs et des compétences
- CNRS : la confiance dans l'expérience des organismes

> Carte des UMR

2. Domaine(s) de recherche

- un pivot : les corpus oraux
- six programmes structurants
- deux programmes de spécialité
- quatre thématiques transversales

> Organigramme de la recherche

3. Prospective

4. Organisation administrative et financière

- administration
- financement et budget

> Organigramme administratif

5. Indicateur des objectifs et échéancier (2012-2015)

- Structuration du laboratoire : intégration et encadrement
- Partenariats : réseau interrégional, réseau thématique, réseau international
- Investissements et financements
- Valorisation
- Diffusion

RESUME

Le Laboratoire Ligérien de Linguistique (LLL) a accumulé au cours du précédent contrat une expertise sur le traitement des corpus oraux. Le travail effectué au sein des équipes a mis l'accent sur la constitution et le traitement des données orales à partir de plusieurs langues (français, anglais, espagnol, créoles à base portugaise, ikwéré, dagara, sèmè, palikur, chiac...). La recherche est allée d'une réflexion d'ensemble sur les principes de collecte (y compris les aspects juridiques et techniques), sur leur signification sociale et sur les instruments de transcription jusqu'à certaines formes d'exploitation des enregistrements et des bases de données concernant des domaines tels que la dynamique diachronique, le rapport écrit/oral ou l'interprétation sémantique de la prosodie. La communauté d'intérêt pour une linguistique de corpus fondée sur l'oral a permis l'émergence d'un travail collectif sur des langues différentes et avec des approches variées. L'intérêt porté à la conservation et à l'exploitation du document sonore, les interrogations sur la représentativité des échantillons, le catalogage et le codage des sources, l'accessibilité et l'interopérabilité ont conduit à un rapprochement avec les personnels du Département de l'Audiovisuel de la Bibliothèque nationale de France. Dans la continuité des travaux conduits dans chacune des institutions, et dans le respect de leur spécificité, la réunion des compétences du LLL et de la BnF permettra de déployer l'ensemble de la chaîne des opérations de saisie et d'apporter au LLL les compétences qui lui font défaut en matière d'archivage, de sauvegarde et de communication, le LLL contribuant de son côté par sa connaissance des procédures de constitution des données. L'UMR constituée sous l'égide de l'INSHS du CNRS en partenariat avec le Ministère de la Culture et de la Communication s'intégrera dans le projet de constitution d'un corpus de référence du français oral préconisé par l'ILF et dans la recherche internationale avec les institutions correspondantes pour d'autres langues (e.g. le BNC – British National Corpus anglais).

SUMMARY

During the previous contract, the Loire Linguistics Laboratory (LLL) has built up expertise in treatment of oral corpora. Teams have focused on the setting up and treatment of oral data from a variety of languages (English, French, Spanish, Portuguese-based creole, Ikwere, Dagara, Sow, Palikur, Chiac ...). Research ranged from a comprehensive reflection on the principles of collection (including legal and technical aspects), their social significance and transcription tools to certain forms of exploitation of records and data within different areas such as dynamic diachronicity, written / oral reports or the semantic interpretation of prosody. The community of those interested in oral corpus enabled the emergence of team work on different languages and with different approaches. The interest in conservation and exploitation of sound recording, the questioning of the representativeness of samples, resource cataloging and coding, accessibility and interoperability have led to a joint venture with researchers and engineers of the Audiovisual Department of the National Library of France (BnF). Continuing what is being carried out within each institution, and with respect to their specificity, combining the skills of the LLL and those of the BnF will allow to span the entire chain of operations involved in the capturing of data and will bring the LLL the skills it lacks in terms of archiving, backup and communication, while LLL will contribute with its knowledge of procedures for setting up data. The UMR (University research team) formed under the auspices of the CNRS INSHS (Social Sciences Institute) in partnership with the Ministry of Culture and Communication will take part in a the projected setting up of a reference corpus of oral French advocated by the ILF (Institute of French Linguistics) and also in the international research involving equivalent institutions for other languages (eg the BNC - British National Corpus).

Le Laboratoire Ligérien de Linguistique (LLL) a été constitué en EA en 2008. Il réunit une trentaine de linguistes (SDL, anglicistes, germaniste, hispaniste) des universités d'Orléans et de Tours et un DR CNRS en détachement en Guyane. Le travail réalisé au cours du précédent contrat a incité le LLL à renouveler son inscription institutionnelle à l'échelle du PRES Centre Val-de-Loire et, en partenariat avec le Département de l'Audiovisuel (comprenant l'ancienne Phonothèque Nationale) de la Bibliothèque nationale de France, à présenter une demande d'association avec le CNRS sous forme d'UMR sur un thème qui correspond aux programmes et aux compétences de chacun des partenaires : les corpus oraux concernant notamment, mais non exclusivement, la langue française. Atteindre cet objectif peut modifier en partie le contenu de ce dossier et jusqu'au nom du laboratoire...

1. PRESENTATION DE L'UMR LLL/BNF/CNRS

L'initiative de transformation d'une Equipe d'Accueil en UMR résulte de la convergence de travaux effectués dans le laboratoire et des études conduites à l'intérieur du Département des documents audiovisuels et sonores de la BnF. Les discussions conduites entre les deux institutions, à travers la Délégation Générale à la Langue Française et aux Langues de France en particulier, ont mis en évidence la nécessité de donner à des collaborations ponctuelles une traduction institutionnelle comme il en existe déjà, au Ministère de la Culture et de la Communication, avec :

- le Centre de Recherches sur la Conservation des documents graphiques – CRCDG (avec le Muséum National d'Histoire Naturelle),
- l'Institut de Recherche sur le Patrimoine Musical en France – IRPMF, le seul impliquant ès qualités la BnF,
- le Centre de Recherche et de Restauration des Musées de France – C2RMF

Le LLL réunit des enseignants-chercheurs de linguistique, quelle que soit leur langue de spécialité, intervenant dans les départements de sciences du langage, d'anglais, d'allemand et d'espagnol des UFR, à l'IUFM ou en IUT des universités de Tours et d'Orléans. Au 1er janvier 2011, dans les deux universités du PRES, se sont prononcés pour appartenir au laboratoire au cours du prochain quadriennal :

- 6 professeurs (+ 1 professeur émérite) de 7^e et 11^e sections, 4 à Orléans et 2 à Tours,
 - 22 MCF, 16 à Orléans (dont 3 à l'IUFM) et 6 à Tours
- (+ en attente, un poste MCF Sciences de l'éducation / SDL demandé à Orléans).

De vingt à trente doctorants sont inscrits dans le laboratoire selon les flux annuels. On compte 2 PR et 1 MCF d'informatique à Tours, 1 MCF de psychologie et 1 de didactique à Orléans en rattachement secondaire, ainsi qu'une cinquantaine de collaborateurs, enseignants du primaire et du secondaire, techniciens de la BnF...

Au-delà de la région Centre, le LLL a conduit des projets en partenariat avec Limoges et Poitiers dans le cadre d'une fédération (FELINCO, 2008-2011) et de conventions bilatérales. L'intégration des deux universités dans le CERLICO (Cercle de Recherche Linguistique du Centre-Ouest) esquisse une forme d'organisation plus large, et plus souple, à l'échelle interrégionale.

Dans ce projet, la Bibliothèque Nationale de France apporte son expertise aux chercheurs en matière de conservation et de mise à disposition tout en valorisant ses collections. Fort d'une expérience qui a fait de cet établissement la référence en matière de conservation, de restauration et d'archivage, le Département de l'Audiovisuel souhaite développer une interaction entre analyse des contenus et problématiques de conservation pérenne à long terme et mise à disposition auprès du public via des outils adaptés : il souhaite ainsi accroître par la recherche la valorisation de ses collections. Une collaboration avec le LLL sur les corpus oraux a semblé, avec les relations de confiance qui se sont nouées au cours des années passées sur des thématiques tels que le *Guide des corpus oraux* ou la préparation du centenaire des « Archives de la Parole », une ouverture pertinente et mutuellement

profitable, annonçant un élargissement à terme vers l'ensemble des laboratoires impliqués dans le domaine selon une structure qui reste à définir.

Postulant au statut d'UMR, le LLL a déjà assuré :

- son intégration dans une fédération de recherche, l'Institut de Linguistique Française (ILF, FR 2393) où il était initialement la seule EA auprès d'une douzaine d'UMR,
- son insertion dans la Maison des Sciences de l'Homme de Tours et dans ce qui en sera le répondant à Orléans,
- sa présence comme membre fondateur lors de la création du CRDO (Centre de Ressources pour la Description de l'Oral) à Aix-en-Provence et Paris,
- l'accueil d'une DR du CNRS les années passées et celle d'un CR pour le prochain contrat,
- sa représentation par un élu dans la section 34 de l'INSHS.

Pour la section 34 du CNRS, affectée par une rétraction du nombre des laboratoires qualifiés après la perte de la qualité d'UMR des équipes de Rouen, Caen, La Réunion et de « Langue Musique Société » sur le campus de Villejuif, cette création renforcerait le potentiel de la section en apportant la preuve de son dynamisme. En même temps, elle contribuerait à une extension de la couverture CNRS des sciences du langage.

Note sur l'implantation des laboratoires CNRS en sciences du langage

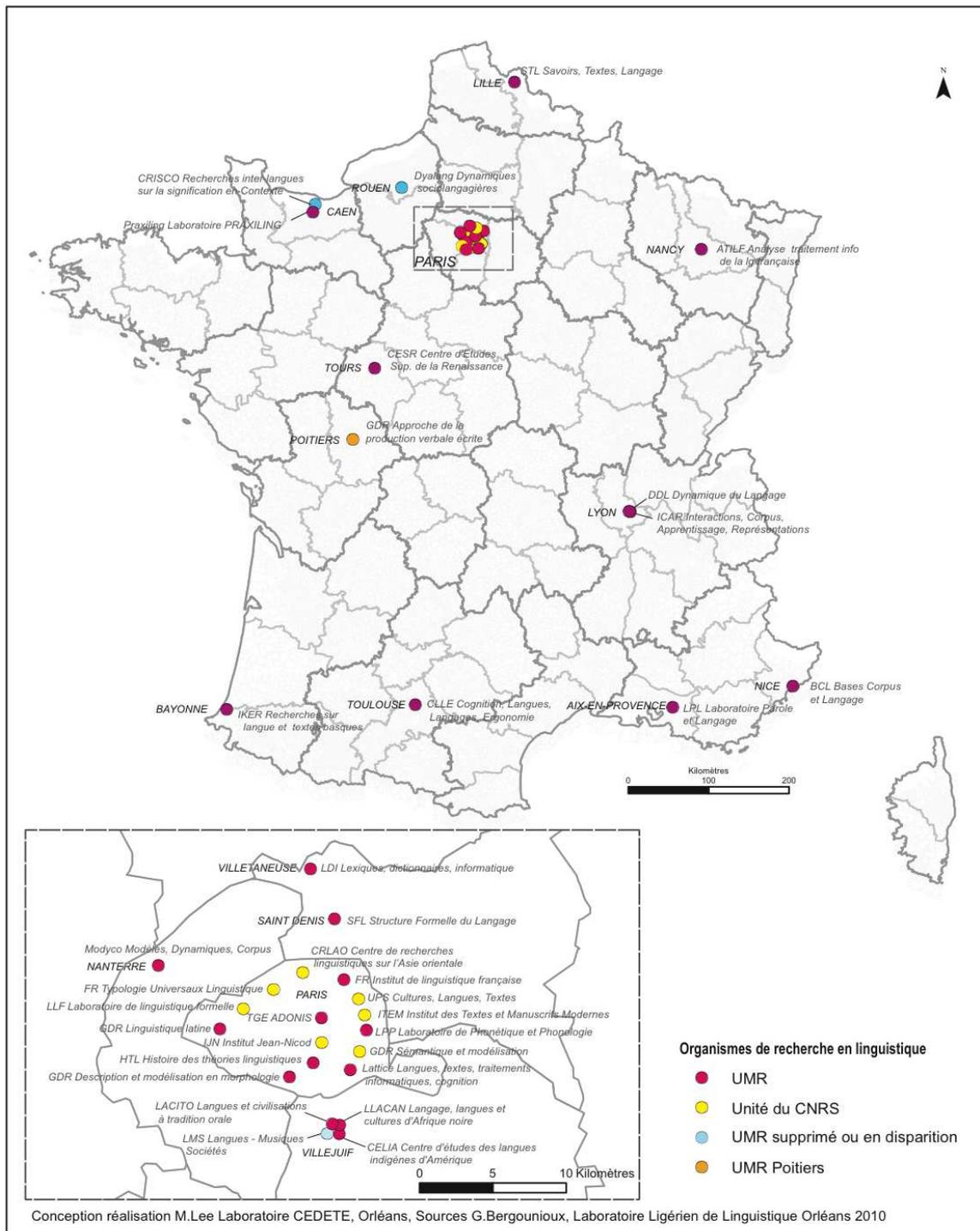
Sur 18 laboratoires dont le rattachement principal relève de la section 34 :

- 10 sont en Ile-de-France, comme les deux fédérations de recherche
- 8 en province (2 à Lyon, 1 à Aix-en-Provence, Nice, Montpellier, Toulouse, Bayonne et Nancy)
auxquels il faut ajouter le laboratoire de Lille rattaché à la section 35. Les effectifs de CR et DR affectés à ces laboratoires accusent le déséquilibre plus qu'ils ne le réduisent.

La carte des implantations montre que le maillage des sciences du langage n'est plus assuré aujourd'hui sur toute la partie Ouest et Centre du pays, au préjudice de la discipline exclue de toute labellisation CNRS à travers un continuum de régions représentant 22 universités, de Brest à Dijon et Besançon et d'Amiens à La Rochelle et Clermont-Ferrand.

CARTE DES EQUIPES DE RECHERCHE CNRS

Implantation des organismes de recherche en linguistique en France métropolitaine



2. DOMAINE(S) DE RECHERCHE

On distinguera **cinq niveaux d'implication** dans les thématiques développées par le laboratoire :

- **un pivot** destiné, par la contribution de tous les chercheurs, à placer le LLL à un niveau d'excellence par la spécialisation sur quelques segments du **processus de collecte et de traitement des corpus oraux**, en lien avec la BnF et en association avec les autres laboratoires et institutions potentiellement concernés, en France et à l'étranger,
- **six programmes structurants** d'accompagnement, associant de trois à huit enseignants-chercheurs du laboratoire,
- **deux programmes de spécialité**, en soutien des formations, engageant de 1 à 2 EC,
- **quatre thématiques transversales** et
- **les recherches individuelles** qui contribuent au rayonnement du laboratoire et à sa présence dans des champs variés (théorie de la traduction, acquisition scripturale, critique littéraire, musicologie...) sans que les productions ne s'intègrent à un projet conditionné par les thématiques (Ce point est explicité dans les fiches individuelles des chercheurs).

2.1 Un pivot : les corpus oraux

Si les corpus sonores ont émergé de facto il y a une centaine d'années comme de possibles moyens de recherche à la disposition des sciences humaines et sociales (Bergounioux, 1992), il a fallu attendre la diffusion du magnétophone pour qu'il soit possible de recueillir et de conserver facilement de la parole sur le terrain. Il en est résulté dès lors un accroissement exponentiel des enregistrements disponibles tel que le recours à des moyens informatiques s'est imposé afin d'en rationaliser la conservation et l'exploitation. Ainsi, en une vingtaine d'années, la référence à des sources contrôlables est devenue inhérente à l'argumentaire scientifique des linguistes mais aussi des anthropologues, des sociologues... et la collaboration avec les informaticiens une nécessité de la recherche.

C'est à cette pratique, sinon nouvelle du moins récente, qui conditionne largement le travail dans les laboratoires, qu'entend se consacrer le LLL avec une double spécificité :

- d'**objet** : les investigations sont centrées sur la classe particulière des corpus langagiers préservés dans leur forme sonore après qu'ils ont été établis ou collectés à des fins d'étude et de transmission, avec l'objectif de rendre compte de leurs propriétés, techniques et documentaires, sans obérer leur conservation ;
- de **méthode**, en unifiant des appropriations fragmentaires par une maîtrise de l'ensemble du processus qui aboutit à la constitution d'un corpus, de sa conception à sa consultation, avec une attention particulière concernant les procédures d'archivage.

Sans remettre en cause le fonctionnement en équipes sur des aires linguistiques ou sur des domaines de la linguistique théoriques (sémantique, phonologie, morphologie, syntaxe) ou appliqués (didactique, TAL), le LLL a en priorité défini pour objectif la maîtrise des corpus oraux, de la conception de l'enquête à la mise à disposition du public et aux implications dans d'autres champs.

2.1.1 Philosophie du projet

Bien que l'étude soit volontairement restreinte aux documents sonores, trois types d'éléments sont impliqués dans l'inventaire :

- les **enregistrements** sous la forme où ils ont été réalisés et/ou conservés, ainsi que les duplications effectuées successivement,
- les **annotations**, qu'elles soient intégrées dans l'enregistrement (sous forme de plages ou de commentaires insérés) ou adjointes sur des supports graphiques au moyen de transcriptions (orthographiques, phonétiques...), d'indications situationnelles et interactionnelles, etc.
- les **outils** et les systèmes de description qui assurent la contextualisation par l'assignation des locuteurs ou de la situation, par l'indexation des contenus, mais aussi qui préservent la maîtrise des techniques de saisie, de gravure et de lecture de l'enregistrement, un corpus sonore étant indissociable des dispositifs qui en assurent la restitution.

Les corpus ainsi entendus correspondent à des produits conçus à l'intérieur d'une pratique scientifique pour être éprouvés par des analyses et pour se prêter à des applications (Habert, Nazarenko, Salem, 1997 *et passim*). Ils peuvent devoir leur existence à un programme qui les a constitués ou bien résulter d'une sélection dans des fonds préexistants. Quelles que soient les sollicitations, les communautés participant à l'exploitation du document sonore ont pour premier souci la préservation du matériel recueilli, ce qui signifie une exploitation non destructive des ressources et l'actualisation continue de l'accès aux contenus afin de surmonter les conséquences de transformations technologiques qui renouvellent formats et appareils de lecture et de compenser la dégradation des supports et des données. La pérennisation des documents suppose la définition d'une méthodologie dont les principes se trouvent définis dans (Baude, 2006).

L'accroissement des réserves de documents sonores et leur prise en charge patrimoniale ont mis en évidence la fragilité des supports et l'obsolescence des systèmes de restitution : les travaux conduits à la BnF assurent une expertise inégalée dans ce domaine. Parallèlement, les fonctions de collecte et de conservation, de stockage et d'analyse, de balisage et d'indexation, de mise à disposition et d'application, effectuées séparément dans chaque laboratoire et centre d'archivage doivent répondre aux exigences induites par l'évolution des technologies et des pratiques scientifiques.

Le document sonore a été appréhendé comme un objet, transférable de l'équipe qui le recueille à l'institution qui le conserve. L'objectif de l'UMR est de concevoir les corpus dans leur dynamique, c'est-à-dire comme un processus, un enchaînement d'opérations solidaires, depuis les hypothèses théoriques sous-jacentes à une décision de collecte jusqu'à la communication au grand public. Dans cette perspective, la collaboration entre les parties contractantes se pense dans une solidarité des gestes scientifiques, une complémentarité des objectifs et des modes de fonctionnement, avec des recoupements et des relais qui confortent la pertinence du projet, sa cohérence et plus encore sa nécessité. C'est une façon de prendre acte qu'une division du travail, qui a prévalu jusqu'à maintenant, ne se justifie plus et qu'il faut participer à la définition de nouvelles formes d'expertise et de connaissance par l'addition des expériences et des compétences. Au-delà, le relais doit être assuré par d'autres structures à l'échelle nationale (TGE ADONIS, CRDO, GDR...) et internationale. Le LLL entend y figurer comme un acteur de premier rang.

2.1.2 Une expertise pour des données spécifiques : les *vocalia*

L'UMR entend contribuer à l'accompagnement scientifique d'une extension des champs documentaires, confinés jusqu'à présent dans les *scripta* (les bibliothèques) ou les *realia* (les musées), en y incluant, dans la spécificité de leur garde et de leur accès public, les *vocalia*. Rapprocher le traitement des corpus sonores des modalités sanctionnées par le travail des équipes qui se consacrent aux recueils de textes et de manuscrits permettra d'achever en région Centre le déploiement d'une expertise qui a déjà fait ses preuves.

Les *scripta*, qui concernent les documents écrits, sont étudiés avec une approche assez semblable à celle que nous entendons déployer :

- par le Centre d'Etudes Supérieures de la Renaissance (CESR / UMR 7576 – Université de Tours) et
- par l'Institut de Recherche sur l'Histoire des Textes (IRHT / UPR 841 – campus CNRS d'Orléans).

Les *realia*, qui appliquent les méthodes des sciences expérimentales à l'analyse des objets à des fins historiques et archéologiques sont représentées sur chacun des sites :

- par l'Institut de Recherche sur les Archéo-Matériaux (IRAMAT / UMR 5060 – Université d'Orléans) et
- par le Centre Interdisciplinaires Cités Territoires Environnement Sociétés (CITER / UMR 6173 – Université de Tours).

On retiendra que les deux laboratoires présents à Orléans, l'IRHT et l'IRAMAT, comptent au nombre des sept laboratoires compris dans la convention passée entre la BnF et le CNRS en 2005.

Les *vocalia* qui concerneront les corpus oraux font appel à des systèmes de représentation et d'indexation pour une définition des procédures d'extraction et de formalisation au moyen desquels peuvent être établis une terminographie, des index, des thesaurus et, aujourd'hui, des ontologies, impliquées dans la gestion des connaissances, les systèmes experts et les procédures d'apprentissage. Quelle que soit l'approche privilégiée, elle demeure tributaire de technologies d'enregistrement et de conservation qui supposent une maîtrise des technologies et des applications. Elle reste engagée dans les représentations qui sont construites à partir des outils de traitement, lesquels déterminent les

modalités de l'indexation. La finalité de la collaboration BnF / Université / CNRS réside dans cette conjonction d'une signalétique, réalisée par l'archiviste, et d'une maîtrise des supports avec une analyse du document effectuée à partir de critères linguistiques.

Les connaissances développées à la BnF recouvrent pour une part les travaux effectués à Tours dans une collaboration étroite entre les chercheurs du Laboratoire d'Informatique et les linguistes du LLL, à Orléans dans la définition d'une convergence entre l'exploration du contenu et la définition signalétique. A partir d'un descripteur formulé au moyen d'un ou plusieurs termes extraits des données documentaires et/ou des transcriptions, sont définis des « pointeurs » qui identifient les séquences du fichier sonore ou le texte correspondant à la recherche formulée. L'application technique conçue en partenariat avec la société ARES en 2010 à partir du corpus ESLO est un exemple de réalisation opérationnelle de ces objectifs. La poursuite de cet objectif, en continuité du travail engagé pour le laboratoire par R. Walter, doit se continuer dans l'insertion des corpus à la BnF et dans les mémoires du TGE ADONIS.

Au-delà du travail, et en lien avec l'*Open Language Archives Community* (OLAC), l'ambition existe de créer, en français et dans d'autres langues, une phonothèque virtuelle mondiale sur le modèle de la bibliothèque virtuelle mondiale. A la suite de la réaction concertée des bibliothèques européennes pour la protection et la mise à disposition de leur patrimoine écrit, c'est à une initiative d'envergure comparable que l'UMR entend apporter sa contribution à travers ses programmes spécifiques. Là encore, la BnF, qui a conduit le projet à l'échelle de l'UE d'une mise à disposition du patrimoine écrit de l'humanité, est le seul organisme public à même de relever un défi du même ordre sur les sources orales de notre mémoire et de notre culture.

Des études préparatoires à la mise à disposition universelle de données traitées selon les standards actuels, l'UMR entend être présente sur un nombre important de segments du « process » (voir tableau suivant) et déterminer les moyens d'une collaboration avec l'ensemble des partenaires actifs dans ce domaine, UMR, EA mais aussi Archives Nationales, directions du patrimoine, associations (AFAS)...

LE PROCESS DANS LES CORPUS ORAUX

Ce tableau est un document interne du LLL. Il a pour finalité d'identifier les compétences et les lacunes du laboratoire dans la chaîne des opérations requises par la constitution d'un corpus.

ETAPES	OBJECTIFS
Préparation de la collecte	Inscription épistémologique du programme et histoire
	Définition des topiques et circonscription des thématiques
	Collation bibliographique et documentaire (état des recherches)
	Définition des complémentarités avec les autres centres de recherche
	Planification et définition des étapes
	Formation des enquêteurs
Détermination du panel et des catégories (échantillonnage et représentativité)	Elaboration des formats d'échange et/ou des questionnaires
	Principes méthodologiques (sociologie de l'interaction, formulations)
	Spécifications des matériels d'enregistrement et contraintes de saisie
	Conduite d'enquête et définition des protocoles
	Rédaction des documents techniques et juridiques, fiches d'enquête
	Recueil des données
Paramètres techniques de la collecte	
Catégorisation des rôles et des situations	
Recueil des métadonnées	
Rédaction d'un journal d'enquête et retour sur expérience	
Documents associés (photos, archive) et modalités de conservation	
Sauvegarde des données	Compatibilité des formats de données polymorphes
	Numérisation pour archivage et réalisation des copies de travail
	Codage des données (définition et pérennisation des formats, TEI)
	Catalogage (principes d'indexation)
	Entretien et renouvellement des supports (restauration, conversion)
	Préparation des transferts (support pivot)
Accès aux données	Enrichissement par des bases de données interopérables
	Conventions de dépôt
	Présence sur Internet et affichage des informations (indexation)
	Mise à disposition ouverte des conventions utilisées
	Adaptation aux requêtes multi-usages
	Définition des seuils d'accès autorisés et des identifiants sollicités
Transcription	Intégration des prescriptions juridiques et techniques
	Alignement de la transcription sur le signal
	Interface vidéo éventuelle
	Interface entre langues (si corpus multilingues ou traduits)
	Assignation des niveaux de transcription (phonétique, morphologie)
	Etiquetage et balises (insertion de commentaires et de gloses)
Traitement des données avant communication	Mode de conservation des différentes versions de transcription
	Protection des données
	Anonymisation
	Intégration d'une boîte à outils
	Indexation et concordanciers
	Définition des formats d'interopérabilité

Recherches linguistiques disciplinaires	Analyse phonétiques (y compris prosodiques) et phonologiques Interactions oral / écrit : représentations, lecture...
	Analyses morpho-syntaxiques et lexicales
	Analyses sémantiques et pragmatiques, confection d'ontologies
	Analyses discursives (stylistiques, rhétoriques)
	Représentations cognitives et constructions mentales
Recherches corrélées et comparatives	Analyses sociolinguistiques et anthropologiques (dont genre)
	Analyses psycholinguistiques, apprentissage et linguistique clinique
	Contact de langues et multilinguisme
	Etudes diachroniques et dynamique
Applications en TAL	Systèmes d'assistance à la transcription
	Conception d'outils pour le traitement de la parole
	Adaptation de BDD (type dictionnaire) aux formats du TAL
	Terminologie (lemmatiseurs) et dictionnaire
	Extraction d'information et structuration de bases de données
	Exploitation statistique et probabiliste
Applications didactiques	Didactique (langue maternelle, langue étrangère et seconde)
	Didactique : langue sur objectifs spécifiques
	Apprentissage de la communication et techniques d'expression
	Pratiques de l'écrit : représentation, acquisition, lecture...
Développements	Conseil en politique linguistique
	Ouverture aux disciplines juridiques
	Inscription de l'oral dans les recherches en sciences sociales
	Contribution aux études artistiques et littéraires
Ingénierie et transfert	Ecriture de scripts d'expérimentation et benchmarking
	Transfert de compétences : assistance à la conception, conseil
	Formation à l'enquête, à la transcription et à la relecture
	Guide d'exploitation des données
Patrimonialisation	Création de phonothèques de laboratoire
	Fonds d'archives sonores (dispositifs de conservation et de lecture)

2.2 Les programmes structurants

Si tous les programmes sont appelés à apporter leur contribution à l'étude des corpus oraux, à des moments différents de leur exécution et de leur traitement, les personnels de la BnF s'y investissant de manière privilégiée dans l'intégration de leurs compétences à une chaîne de traitement, quelques programmes actifs dans le LLL au cours du précédent contrat fourniront une première structuration des orientations.

ESLO

Sémantique énonciative et cognition

Morpho-phonologie

Langues de Guyane

Langues d'Afrique et créoles

Didactique du français et des langues

2.2.1 ESLO

Responsable : O. Baude

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

Doctorants : L. Hriba, S. Moukrim

Autres participants : A. Chesneau, P. Philardeau, M. Petit, J.-L. Rougé

L'Enquête Socio-Linguistique à Orléans (désormais : ESLO 1) conduite en 1968 comprend environ 200 interviews, toutes référencées et au total plus de 300 heures de parole incluant une gamme d'enregistrements variés (conversations téléphoniques, réunions publiques, transactions commerciales, repas de famille, entretiens médico-pédagogiques, etc.). ESLO 1 couvre l'ensemble des catégories socio-professionnelles, hommes et femmes, et présente un échantillon des formats de la communication, des tâches linguistiques et des types de discours selon une approche dialogique. Ce corpus représente, par son ampleur, sa rigueur et sa cohérence, le plus important témoignage disponible sur le français parlé avant 1980 (corpus de 4 500 000 mots environ).

En partant des acquis d'ESLO 1, une nouvelle enquête, dénommée ESLO 2, a été mise en chantier. A quarante années de distance, elle vise à constituer un corpus comparable dans le produit attendu et dans les modalités de la collecte : l'objectif a été fixé à 400 heures environ de documents sonores qui totaliseraient approximativement 6 000 000 de mots. Réunis, ESLO 1 et ESLO 2 formeront une collection de 700 heures d'enregistrement, soit plus de 10 000 000 de mots, ce qui est considéré aujourd'hui comme une valeur repère pour les investigations projetées.

ESLO 2 a été conçu pour préfigurer la référence attendue dans un domaine qui en est encore à se structurer et dans lequel se manifeste de manière récurrente une demande de définition pour un format standardisé de *collecte*, de *conservation*, de *traitement* et d'*analyse* :

- la *collecte* est l'occasion de définir, qualitativement et quantitativement, le profil de l'échantillon représentatif, en particulier dans la sélection des modes d'interactions entre les témoins et les enquêteurs ;
- la *conservation*, qui inclut la préservation des supports, l'indexation des contenus et l'accessibilité (c'est-à-dire la protection) des données, conditionne le partage des sources à des fins d'étude scientifique ou didactique ;
- le *traitement* suppose la maîtrise d'un process qui va de la conversion numérique des enregistrements jusqu'à la réalisation d'une transcription balisée et ouverte ;
- l'*analyse* constitue l'épreuve des théories (et des instruments, en particulier des logiciels) puisqu'elle compare les formalisations et les opérations et qu'elle valide ou infirme les hypothèses en prenant argument de leur compatibilité avec les faits.

La collecte sera close fin 2011 et son dépôt à la BnF constituera, pour les deux parties, un test grandeur nature concernant la qualification et l'exploitabilité du corpus, un retour d'expérience à l'échelle 1.

Par ailleurs, afin de rendre compte de la dynamique du changement, telle que la comparaison diachronique et variationniste des enquêtes le permet, la collecte des données a été étendue à toutes les langues parlées à Orléans. Bénéficiant d'un soutien régional (PANGLOSS), le programme « Langues en Contact à Orléans », piloté par J.-L. Rougé avec l'aide de C. Brumelot, I. Diallo et S. Moukrim prend en compte l'interaction des locuteurs et des systèmes. La première phase concerne un inventaire et une identification des langues parlées avant de situer les conditions dans lesquelles leur exercice peut être observé aussi bien dans des contextes spécifiques mono- ou multilingues que dans l'influence réciproque des langues, incluant des observations sur les enseignes et les affichages ou des entretiens avec les traducteurs assermentés.

L'objectif est d'élaborer une base de corpus oraux fondée sur des matériaux dont la collecte et le traitement pourront soutenir la comparaison avec les exigences apportées par l'INaLF aux corpus écrits :

- (1) en accumulant des heures de parole et en développant une expérience critique sur leur mode de collecte et d'analyse,
- (2) en créant des matériaux exploitables par les chercheurs et en participant à la structuration de la communauté « corpus »,
- (3) en ouvrant des champs d'investigation à l'ensemble des sciences humaines et en apportant les attestations requises par la lexicographie ou l'élaboration de grammaires de référence.

Le projet est aussi destiné à inaugurer une nouvelle politique de conservation du document scientifique sonore non musical et à préparer des collaborations dans les applications, didactiques et informatiques.

Au cours du prochain contrat, le travail engagé par l'équipe sera poursuivi, autant dans une politique suivie d'accroissement des données collectées, par l'extension des enquêtes et la diversification des situations, que par une définition d'une véritable politique d'accès aux corpus oraux. A la différence des réalisations correspondant aux numérisations des bibliothèques, le champ de l'oral reste en friche pour une exploitation en termes de recherche et d'application.

Le caractère sensible des informations et des données, en dépit des précautions prises au moment de la collecte et du traitement pour en assurer l'anonymisation, constituent un frein à l'accès aux corpus. Les répercussions sont sensibles dans une orientation de la linguistique qui privilégie, en sémantique, en syntaxe ou en analyse de discours des données écrites, c'est-à-dire des attestations des langues déjà inscrites dans un format convenu. La mise à disposition sur le portail du Ministère de la Culture de données d'ESLO1, la négociation d'une base de données avec la société ARES ont permis de réfléchir à la façon de présenter les données et de les mettre à disposition avec des accès différenciés selon le type de requête (grand public / chercheur des grands organismes / entreprise / collaborateurs).

Au-delà de la constitution des données, sept objectifs ont été proposés au sein de l'équipe. En premier, et de façon centrale, l'exploitation linguistique des ressources et des attestations. Chaque chercheur a été invité à illustrer l'enrichissement des théories qui lui sont familières et dont il a l'usage en les confrontant aux occurrences disponibles dans l'ensemble du corpus, bénéficiant de la disponibilité offerte par la collecte, les transcriptions et les outils développés sur la base.

Parmi les travaux exécutés, on peut mentionner des études en sociolinguistique de l'interaction et des comparaisons diachroniques, mais aussi une étude phonologique, une recherche en linguistique de l'énonciation, une analyse en TAL des opérations. Une synthèse sera publiée en 2011 autour d'un sous-corpus exemplaire, « l'omelette » (92 recettes d'omelette en réponse à une question spécifique du questionnaire ESLO1), qui associera huit chercheurs de l'équipe afin d'analyser, sous ses différentes facettes, un corpus fourni parallèlement.

Un travail en didactique doit prolonger ces travaux, autant du côté de la collecte de nouvelles données (notamment autour de la question de la transmission de la langue en milieu familial, par construction d'un module qui enregistrerait des enfants et leurs parents afin de mesurer la dynamique interne du changement linguistique. Cette recherche, conduite avec les collègues de l'IUFM qui ont déjà engagé un programme d'enregistrement dans le cadre scolaire, doit être complétée par des applications didactiques, en FLE (comme il avait été prévu initialement pour ESLO1) et en FLM, en collaboration avec le GORDF.

En collaboration avec l'équipe « Créoles et Langue d'Afrique », il est prévu d'exploiter le résultat des travaux conduits au titre de « Langues en Contact à Orléans », en particulier à travers le programme PANGLOSS, sur financement de la Région Centre. La rapidité des changements dans la composition urbaine de la ville permet de mettre en place en temps réel une observation des différentes

communautés linguistiques, de leur solidarités et de leur répartition, de leur présence dans l'école et la vie associative, dans le monde économique aussi, en suivant des trajets migratoires et des recompositions d'usage linguistique. Une thèse aura été soutenue sur ce thème et de nouvelles demandes de contrats doctoraux seront sollicitées, en particulier pour des étudiants d'origine étrangère qui pourront apporter leur connaissance de deux cultures dans le projet.

Ces différentes données sont centrales dans la collaboration avec la BnF et le TGE-ADONIS puisque ce corpus doit servir de test pour les propositions qui seront faites afin d'établir le grand corpus de référence du français. La BnF, par son spectre d'intervention et la richesse de ses collections, a déjà fait la démonstration de ses capacités techniques. Le TGE ADONIS offre une infrastructure numérique ouverte à toutes les communautés scientifiques des sciences humaines et sociales, dans les principaux domaines du numérique : collecte et acquisition, traitements et calculs sur des données, travail collaboratif, publication électronique, hébergement, archivage à long terme des données de la recherche (sources, publications, etc.). Selon la déclaration de ses promoteurs : « L'enjeu majeur est de permettre un accès le plus large possible aux données scientifiques produites par les SHS, tout en améliorant l'interopérabilité qui préfigure la construction du web de données. » A l'image de ce qui a été réalisé sur les archives (IRHT, Paris & Orléans), sur les lexiques et les textes (INaLF, Nancy), sur les sources iconographiques (Paris & Bordeaux), sur l'édition de textes patrimoniaux (CESR, Tours) et en bénéficiant du travail effectué sur la TEI, ESLO entend devenir un partenaire majeur, mais non exclusif, dans le domaine des corpus oraux.

Exploitation en linguistique, définition du corpus de référence du français parlé, applications didactiques mais aussi transfert. L'une des premières initiatives du LLL et de la BnF serait l'organisation d'une école d'été ouverte en particulier aux chercheurs des pays du Sud sur les techniques de numérisation et de conservation des données orales. Une autre attente concerne la réunion des enquêteurs, si nombreux aujourd'hui dans les sciences humaines et sociales, sans qu'une réflexion spécifique sur cette activité en linguistique ait été conduite. Sur le modèle du *Guide des bonnes pratiques / Corpus oraux*, la réalisation de manuels en ligne confiés à des représentants éminents de la communauté serait un outil utile à tous les laboratoires impliqués dans ces démarches.

Enfin, si l'on en juge par la poursuite dans le secteur privé des étudiants en Master formés à Orléans dans le domaine du TAL (un lien avec la formation qui sera continué), par le recrutement d'une doctorante par l'entreprise Sinequa, il y a un créneau disponible pour les compétences développées dans ce domaine en ligne avec les ESLOs.

2.2.2 Sémantique énonciative et cognition

Responsable : G. Col

Participants (EC) : A. Cambourian, S. Gatelais, S. Osu, F. Toupin,

Doctorants : Elisabeth Grimaud, Augustin Ndione

Autres participants : Claire Agafonov, A. Delplanque, T. Gallèpe,

Les objectifs généraux de cet axe sont l'étude de notions ou d'unités linguistiques du point de vue de la construction du sens, avec comme arrière-plan les théories énonciatives françaises et les théories cognitives nord-américaines et européennes.

Tout en constituant un axe de recherche nouveau dans le LLL-Tours, les travaux que nous projetons de mener dans le programme *Sémantique Énonciative et Cognition* sont en fait impulsés par des collaborations antérieures, notamment autour de la question des verbes d'apparence en anglais, allemand, français, et russe (CORELA, 2005).

Cadre méthodologique et théorique

Les théories énonciatives

Nous postulons que le sens associé à un énoncé est un construit résultant de l'interaction des divers constituants de l'énoncé réinterprétés comme marqueurs d'opérations linguistiques (que cela relève des morphèmes, de la prosodie, du placement dans le linéaire, etc.). Nous cherchons ainsi à dégager le mode de fonctionnement spécifique et invariant de ces constituants à travers l'étude minutieuse de leurs emplois les plus variés. On se posera notamment les questions suivantes :

- Quel est le rôle spécifique du marqueur étudié dans la construction du sens de l'énoncé dans lequel il apparaît ?
- Ce rôle spécifique constitue-t-il la « propriété invariante », ou « forme schématique » dans la perspective de la Théorie des Opérations Énonciatives et Prédicatives (Culioli), du marqueur ?
- Quelle place peut être attribuée au signifié de ce marqueur dans la « hiérarchie des connexions » (Fourquet) censée décrire le signifié de l'énoncé ?
- Si l'altérité est de constitution (Culioli), si la prise en compte de l'altérité ne s'effectue pas toujours sur le même mode (Delmas), alors, sur le plan pragmatique, à quelle image de l'autre le marqueur renvoie-t-il ?
- En quoi l'histoire et l'évolution du marqueur nous éclairent-elles sur ses emplois contemporain ? Suivant quelles voies et pourquoi les systèmes se recomposent-ils ?

Le rôle du co-texte et du contexte, ainsi que les conditions (synchroniques, diachroniques) d'apparition de l'unité seront donc particulièrement pris en compte. La prise en compte des structures malformées, voire impossibles, servira par ailleurs de garde-fou et aidera à mieux cerner les rôles spécifiques des marqueurs étudiés.

La problématique cognitive

Sera aussi envisagée une réflexion sur les relations entre le langage et les autres activités cognitives, avec un intérêt particulier pour la perception. Dans cette perspective, les travaux pourront porter sur :

- la description de la construction progressive et dynamique du sens d'un énoncé en cotexte, au fur et à mesure de sa perception. Il s'agira de mettre en place et de valider une formalisation dynamique issue notamment des travaux en linguistique cognitive (Fauconnier, Talmy), en grammaire cognitive (Langacker) et en psycholinguistique cognitive (Chafe, Tomasello, Le Ny).

Une approche si possible pluridisciplinaire (psychologie, informatique, système dynamique) en relation avec les sciences de la cognition de manière plus générale sera recherchée.

- la définition d'une approche « instructionnelle » de l'analyse des unités linguistiques : quelles instructions sémantiques sont fournies par les unités d'un énoncé ? Comment définir le terme « instruction » ainsi que son statut en linguistique ? Comment mettre en évidence ces instructions ? Comment décrire l'assemblage des unités par les instructions qu'elles fournissent ?
- enfin les notions de « forme schématique » de la TOPE et celle d'« instruction » développée dans le cadre cognitif seront comparées finement afin de mettre en évidence la spécificité de l'un et de l'autre, ainsi que leur complémentarité.

Loin d'être antinomiques, les deux approches suggérées pour mener à bien ce projet de sémantique énonciative et cognitive sont effectivement complémentaires et ceci sur au moins trois points essentiels :

- l'approche contrastive qui sera mise en place
- l'analyse à granularité différente qu'elles impliquent l'une et l'autre (morphème, unité grammaticale, syntagme, « construction »)
- la définition d'un paradigme complet, intralinguistique et interlinguistique (étude comparative d'unités proches), ainsi que la relation du langage aux autres processus cognitifs.

Programmes et thèmes de travail

Opération 1 : Analyse et formalisation de l'expression du manque.

La thématique générale de notre travail portera sur l'analyse et la formalisation de l'expression du manque. Nous tenterons d'aborder cette thématique à la fois du point de vue des formes linguistiques elles-mêmes, mais aussi du point de vue de la conceptualisation de la notion de manque. Les questions que nous serons amenés à traiter concernent différents domaines des langues représentées dans l'axe (allemand, anglais (contemporain et médiéval), dagara, français, ikwéré, russe). Suivant ces langues, certaines questions se révèlent plus pertinentes que d'autres. Nous suggérons ainsi une série de problématiques à aborder, de manière non exclusive, suivant différentes orientations :

Orientation syntaxique :

- La question de la diathèse, par exemple à travers l'opposition « tu me manques » vs. « I miss you »,
- La question de la rection (transitivité vs. intransitivité) à travers des exemples de construction comme celles du verbe « manquer » (*manquer de (pain), manquer une cible, manquer tomber, etc*), ou celles du verbe « fail » (*fail to do something, fail in one's duty, fail a test, etc*).

Orientation morphologique :

- La question des affixes évoquant le manque, comme le préfixe *pro-* en russe, *-less / -los* en anglais et allemand. Nous nous demanderons par ailleurs s'il existe des prépositions ou des particules indiquant le manque (on pense notamment à des unités comme *court* dans « être à court de », ou *out of, away*), et sinon, pourquoi.
- Certains verbes russes comme *ne xvatat'* et *ne dostavat'* se construisent avec le génitif de l'entité absente (ou en quantité insuffisante), et éventuellement le datif de la personne concernée, ou un complément de lieu. D'où des questions comme : quand préfère-t-on un datif vs un locatif ? (*cf* en français *il manque une page dans ce livre / à ce livre*) ?

Orientation lexicale :

- Le domaine de la sémantique lexicale ne sera pas limité à la question du verbe, mais s'ouvrira sur des unités comme : *need, lack, fail, manquer, rater, le manque, raté, want* (nom/verbe), *lack, failed*.

Orientation énonciative :

- Dans le domaine nominal plus particulièrement, on peut aborder la question de la *restriction*, c'est-à-dire du manque par rapport à une attente (*little* opposé à *a little* ; *peu* / *un peu* ; *wenig* ; etc.)
- La langue russe introduit par ailleurs la négation pour exprimer le manque, par exemple à travers les deux verbes cités plus haut signifiant « manquer » au sens d'« être absent » ou bien, suivant le contexte, « être en quantité insuffisante », « ne pas suffire » : *ne xvatat'* et *ne dostavat'*. Qu'observe-t-on dans les autres langues ? Quel est le rôle de la négation dans la perspective de l'expression du manque ?

Orientation cognitive :

- Est-ce que les différentes langues étudiées nous permettent de mettre en évidence une conceptualisation commune de la notion de manque ? Si oui, quelle représentation proposer pour cette conceptualisation et comment formaliser cette représentation cognitive de la notion de manque ?
- Est-ce qu'au contraire, l'analyse linguistique nous permet de mettre en évidence des conceptualisations différentes de la notion de manque ? Il s'agira alors d'identifier les raisons de ces différences en décrivant finement ces conceptualisations, et en proposant une formalisation appropriée.

Opération 2 : sémantique instructionnelle et prosodie

La sémantique linguistique a comme objets principaux de rendre compte de la diversité des emplois des signes linguistique d'une part (polysémie) et d'autre part de rendre compte du processus d'intégration sémantique entre ces signes, et ce du morphème au discours.

La première question a vu émerger une sémantique instructionnelle capable de rendre compte de la façon dont un même input sémantique (instruction) conduisait dans ses différents emplois à des interprétations diverses (sens) qui pouvaient se lexicaliser.

Depuis quelques années, l'étude de la diversité des emplois a commencé à être menée en prenant en compte la dimension prosodique des emplois, autrement dit le fait que selon la façon dont il est réalisé (ou selon le contexte prosodique), un signe va recevoir des interprétations distinctes et parfois opposées. Sans rentrer dans tous les aspects de la question, il ressort de ces premiers travaux, qui ont permis de définir et de valider une méthodologie associant analyse sémantique et analyse prosodique (Petit, 2009), que l'intonation contraint effectivement l'interprétation des emplois (et génère donc de la polysémie) et qu'il est possible de cartographier prosodiquement les emplois d'un signe et en généralisant, l'ensemble du lexique.

C'est dans ce contexte que se situe le programme de l'opération SIP (« Sémantique Instructionnelle et Prosodie ») qui associe les deux équipes du LLL, ainsi que des chercheurs du LaTice (voir « Collaborations nationales »).

Dans son versant tourangeau et concernant l'anglais, l'objectif est d'analyser et de formaliser l'instruction sémantique fournie par des unités discursives de très haute fréquence de l'anglais comme *well*, *still*, *always*, *actually*, *yes*, *no* ou d'autres unités de ce type (sur le modèle des analyses proposées dans Col 2009 et Col 2010). La définition des instructions fournies par ces unités se fera par ailleurs en associant étroitement les phénomènes intonodiscursifs associés à leurs emplois. Il s'agira donc d'établir tout d'abord un banque d'emplois interrogeable de ces unités afin de les caractériser sémantiquement (définir leur « forme schématique », Culioli 1990), puis de combiner leur définition sémantique avec leur comportement prosodique. Les unités choisies sont soit des unités à potentiel holophrastique (*well*, *actually*), ou bien au contraire des unités fortement dépendantes qui introduisent un paquet d'unités (*still*,

always, yes, no). Pourront également être comparées avec les unités citées plus haut d'autres unités comme des prépositions ou bien des conjonctions de subordination afin d'apporter un contre-point aux analyses. Ces dernières seront menées sur des corpus déjà constitués comme le corpus COLT (Corpus of London Teenagers) ou le corpus Aix-MARSEC du laboratoire LPL (CNRS/Université de Provence), mais aussi sur des corpus plus ciblés sur le plan discursif. Afin de travailler sur la détection des prééminences et des faits intonatifs saillants, le logiciel ANALOR développé au LaTTiCE par B. Victorri sera utilisé et paramétré pour l'anglais.

Collaborations

Université d'Orléans

Pierre Cadiot (PR/LLL Orléans), François Nemo (PR/LLL Orléans), Noëlle Serpollet (MCF/LLL Orléans), Mélanie Petit (ATER/LLL Orléans) dans le cadre de l'opération « Sémantique instructionnelle et prosodie ».

CNRS

LaTTiCe, UMR 8094, ENS-CNRS (Jeanne Aptekman (post doc), Thierry Poibeau (CR CNRS), Stéphanie Girault (IR CNRS), Bernard Victorri (DR CNRS)), dans le cadre de l'opération « Sémantique instructionnelle et prosodie ».

CORPUS (Amiens), EA 4295, qui co-organise avec LLL-Tours le Colloque Bisannuel de Diachronie de l'Anglais, (prochain colloque à Tours prévu en 2011).

Centre de Recherche en Linguistique Interlangues et Traitement des Textes de Poitiers (CeRLITeP) du laboratoire FoReLL (EA 3816 / Université de Poitiers) : Jean-Louis Duchet (PR), Sylvie Hanote (MCF), Nicolas Videau (doctorant) dans le cadre de l'opération « Sémantique instructionnelle et prosodie ».

2.2.3 Morphophonologie

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. 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^e 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 poursuite de cette activité étant dans la continuité de celle conduite dans le contrat en cours, on se reportera au bilan pour le détail de la méthode et des études, à la bibliographie des participants pour les publications.

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) auront permis la réalisation de la base. L'interfaçage avec la base de données historique constituée à Poitiers doit permettre 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^e s.), afin d'établir une édition numérique du dictionnaire de John Walker, *A Critical Pronouncing Dictionary And Expositor of the English Language* (2^e éd, 1797), qui cumule les observations de nombreux autres auteurs du XVIII^e 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 (2012-2015).

Ce projet, qui s'intègre dans la relance, par les universités de Tours et d'Orléans, du Réseau Français de Phonologie, a fait l'objet d'un financement des Universités de Tours et de Poitiers pour sa phase initiale.

2.2.4 Langues de Guyane

Responsable : A. Cristinoi & F. Grenand

Participants (EC) : P. Cadiot, F. Nemo

Collaborations : S. Bahuchet, G. Bergounioux, A. Capiberibe, G. Collomb, P. Grenand, D. Tilkin-Gallois

Ce programme a été entrepris, en partenariat avec le Ministère de l'Outre-Mer et le Ministère de la Culture, pour réaliser la description, linguistique et anthropologique, de trois nations guyanaises : les Palikur de langue arawak, les Wayãpi de langue tupi-guarani et les Wayana de langue karib.

L'une des originalités de ce projet tient dans l'articulation entre la réalisation d'un travail linguistique, la place centrale accordée à la dimension sociale de la langue et la restitution d'une culture, dans son intégralité passée et présente, par la réalisation d'une encyclopédie. Chacune des quatre séries encyclopédiques (une pour chacun des peuples plus une pour les questions transverses) a vocation à être une suite de fascicules thématiques traitant de la vie de ces communautés avec un accent particulier mis sur les aspects linguistiques

Pour ces encyclopédies se trouvent combinés deux principes, l'un monographique (progression thématique), l'autre dictionnaire (progression alphabétique). Le concept est inédit en ce sens que certains fascicules se prêtent à la forme classique des articles, d'autres à la forme de la monographie. Ils paraissent au fur et à mesure de leur achèvement, en donnant priorité à certains thèmes majeurs. Ainsi ont déjà paru une étude sur l'histoire orale des Wayampi (J. Chapuis), une recension ethnographique des Palikur (Nimuendaju) et la présentation générique des trois collections (P. et F. Grenand).

L'ensemble est destiné à accueillir, en suivant l'annonce faite dans la présentation :

- des études thématiques,
- une bibliothèque rassemblant les textes de littérature orale produit par chacun des trois peuples et les textes d'archives produits par ceux qui les ont visités,
- un dictionnaire fondamental extrait d'un grand dictionnaire électronique de langue,
- une grammaire, avec possibilité d'applications didactiques.

Ces peuples ayant été longtemps tenus en marge de l'écriture, on mesure l'apport de ce projet à l'étude des corpus oraux. L'objectif, au cours du prochain contrat, est de publier, au rythme d'au moins un et si possible deux fascicules par an les volumes de ces encyclopédies tout en préparant une version en accès public d'un dictionnaire du palikur comprenant au moins 5000 entrées, 3500 de vocabulaire de base et 1500 d'un lexique botanique. La première étape (en 2011) sera la publication d'une phonologie du palikur fixant les principes de transcription et l'édition des fiches botaniques déjà préparées.

2.2.5 Langues d'Afrique et créoles

Responsable : J.-L. Rougé

Participants (EC) : G. Fabre, S. Osu, E. Schang

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

Collaborations : R. Boyd, A. Delplanque

Ce programme concerne extensivement une demi-douzaine de pays d'Afrique subsaharienne (Côte d'Ivoire, Centrafrique, Cameroun, Nigéria) avec une insistance particulière d'une part sur le Burkina-Faso, en collaboration avec le CNRST à Ouagadougou, d'autre part sur les pays où sont parlés des créoles à base portugaise (Cap-Vert, Sénégal, Guinée-Bissau, Sao-Tome et Principe)

Pour les créoles à base portugaise, le LLL est devenu le premier centre mondial en ce domaine puisqu'il est le seul qui ait deux chercheurs spécialisés dans ce domaine ! Au-delà de l'anecdote qui montre néanmoins la nécessité de structurer ce domaine, on notera d'un côté le recouvrement du domaine puisque le dictionnaire de référence a été publié par l'un des deux chercheurs qui en prépare une version électronique tandis que l'autre s'est signalé par sa capacité à appliquer les méthodes et les théories les plus actuelles sur un champ qui apparaît souvent en friche (organisation de journées d'études avec le LLF, le laboratoire SFL et des collègues étrangers). Il est attendu, au cours du prochain contrat :

- un ouvrage de synthèse à paraître dans la collection « Les Langues du Monde » de la Société de Linguistique de Paris, complété éventuellement par une monographie,
- le traitement des corpus recueillis depuis des années pour extraction automatisée d'informations,
- l'organisation d'une phonothèque de ces langues,
- l'étude dialectologique des variétés dans les communautés émigrées en France,
- la numérisation du dictionnaire des créoles à base portugaise paru aux éditions Karthala.

Ces recherches s'appuient sur un accompagnement au développement des infrastructures de l'enseignement supérieur dans les pays concernés, en matière de linguistique, d'enseignement du français et de formation des maîtres.

Au Burkina, la signature d'une convention en appui en matière de formation et de recherche permettra d'assurer un fort développement des échanges. La partie formation comprend des enseignements de Master destinés aux étudiants de l'université de Ouagadougou, l'accueil de doctorants (deux actuellement) et des formations sur la préservation des corpus oraux et leur exploitation (un stage en 2007, un en 2010). La recherche, en complémentarité du travail mené par nos collègues burkinabès sur la cinquantaine de langues vernaculaires, se développe autour de la description et de l'analyse de deux langues : le dagara et le sèmè dans le cadre du programme Radicel-K. Il s'agit de parvenir à doter ces deux langues des outils qui en assurent la recevabilité dans la communauté scientifique en même temps que sera assurée une appropriation lettrée par les locuteurs. Au terme du prochain contrat, est attendu un ouvrage sur le dagara publié dans la collection « Les Langues du Monde » par A. Delplanque et un ensemble de contributions, en collaboration avec les chercheurs du pays, sur la description des langues nationales.

Un volet original de ce projet est constitué par le développement des TIC au service de la collaboration entre chercheurs, de la diffusion des matériaux et de la valorisation des résultats grâce à la mise en œuvre de méthodes et de programmes informatiques originaux permettant l'exploitation collective de bases de données pluridisciplinaires évolutives élaborées grâce à différents outils dédiés à la linguistique. Est prévues la mise en place d'un environnement numérique de travail collaboratif entre cinq partenaires : les deux universités de la région Centre, l'IRD, le CNRST et l'Université de Ouagadougou,

afin de tester concrètement un dispositif qui permette la mise en ligne, l'archivage et le partage des données sonores et transcrites et la diffusion des résultats. Un ingénieur informaticien est associé à l'équipe de linguistes afin d'assurer la formation des partenaires burkinabé aux outils dédiés à la linguistique descriptive (une mission annuelle les deux premières années).

Outre la possibilité d'avancer plus rapidement dans notre projet grâce à l'accessibilité aux données engendrée par cette plateforme, ce dispositif permettra d'initier une pratique innovante de recherche collaborative à distance. Le dispositif technique envisagé devra, au besoin, être adaptable à d'autres collaborations, tant à l'intérieur du projet (collaborations entre et avec les géographes et écologues) qu'à l'extérieur du projet puisque le champ d'action du groupe de recherche sur les langues africaines du LLL, de par les compétences spécifiques de ses différents participants, couvre de nombreux autres pays.

S'agissant de la mise en place d'une collaboration entre des instituts de la Région Centre et des partenaires d'Afrique subsaharienne, qui rencontrent de fréquentes difficultés pour trouver des fonds, le financement public de cet aspect du projet est une nécessité pour réduire la fracture numérique existante. Aussi, la mise en place d'un protocole efficace de collaboration effective entre différents centres de recherche de pays du Nord et du Sud situera les universités de la région Centre dans l'innovation pour la collaboration numérique et le transfert de technologie en général avec les pays émergents.

Parmi les autres initiatives, la poursuite du travail sur l'ikwéré et une étude sur le kota du Gabon à partir d'un de ses dialectes.

2.2.6 Didactique du français et des langues

Responsable : J. Lafont-Terranova

Participants (EC) : Ph. Bourdier, F. Wolf-Mandroux

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

Collaborations : D. Colin, S. Chevillard, A.-L. Doyen, G. Guetemme, J. Isidore-Prigent, D. Ulma

La première ambition est de penser le passage de l'ordre de l'oral à l'ordre du scriptural. Les recherches en didactique de l'écriture sont centrées sur les conditions d'acculturation à l'écrit qui favorise la conceptualisation nécessaire pour la construction de savoirs qui supposent la maîtrise d'un usage normé de la langue.

Le Groupe Orléanais de Recherche en Didactique du Français (GORDF) a pour projet de développer, en collaboration avec les collègues du LLL, une réflexion sur le recueil des données et la constitution des corpus : détermination du panel et des catégories dans une perspective qualitative, élaboration des formats d'échange et/ou des questionnaires, établissement des protocoles de recueil, formation des enquêteurs aux principes méthodologiques de recueil de données (ex. conduite d'entretien semi-directif), collecte, transcription, mise à disposition des corpus... Il s'agira notamment de formaliser les conditions d'adéquation entre les objectifs de recherche en didactique et les conventions retenues pour la transcription, concernant en particulier le degré d'« oralité » à conserver. Cette réflexion sera élargie au recueil et au traitement de données audio-visuelles.

Une partie du travail se déroulera en partenariat avec les chercheurs de l'IUFM Centre Val-de-Loire et le laboratoire de psychologie de l'université de Poitiers, dans le cadre des programmes spécifiques déposés par la MSHS. Une réflexion est engagée pour une exploitation didactique de l'enquête ESLO, en partenariat avec la spécialité « Linguistique et didactique » Master Linguistique, qui a des prolongements en anglais dans une collaboration avec le CIEP.

Recherches

Participation au projet « Les conditions de maîtrise de la production écrite » (resp. E. Lambert, MCF psychologie, CeRCA), programme déposé par la MSHS de Poitiers

Deux études sont prévues pour poursuivre le travail amorcé avec le Plan Pluri-Formations *Ecole, Education et Sociabilités* (2008-2011).

La première étude se situe dans l'axe 1 du projet (« La gestion cognitive des différents niveaux de traitement au cours de l'acquisition de la production écrite »). Cette étude s'appuiera sur deux terrains d'expérimentation visant chez des jeunes apprenants d'une part et chez des lycéens et étudiants d'autre part, à mieux connaître certaines conditions de production d'écrit.

Expérience 1 : Le projet s'appuiera sur une approche développementale des pré-requis (moteurs et procéduraux) à la production de mots dans une population d'élèves de 4 à 7 ans et s'inscrira dans une visée pédagogique. Cette recherche se fera en collaboration avec Magali Noyer-Martin (MC psychologie).

Expérience 2 : Utilisation des abréviations chez des lycéens et étudiants dyslexiques-dysorthographiques. L'objectif est de montrer que les abréviations sont sous-utilisées par les élèves et étudiants dyslexiques-dysorthographiques et de caractériser le rôle des compétences langagières qui en permettraient une utilisation efficace.

La seconde étude se situe dans l'axe 3 du projet (« Transitions scolaires, transitions professionnelles et pratiques différenciées de l'écrit »). L'étude repose sur le constat que le passage du primaire au collège demande à l'élève un repositionnement important en matière de pratique scripturale. Il s'agira d'approfondir l'analyse du rapport à l'écriture des élèves et des enseignants des deux niveaux à partir de l'exploitation des données recueillies en 2008-2010. L'étude portera notamment sur les transcriptions d'entretiens menés en 2008-2010 et se focalisera sur deux disciplines, le français et les sciences. Dans une perspective curriculaire, les données recueillies seront confrontées aux prescriptions institutionnelles de 2002 et de 2008. Il est prévu de déposer un projet de doctorat lié à cette étude.

Un pan de la recherche pourra s'associer au groupe EPISTEVERB (IUFM de Lyon, Grenoble, Aix-Marseille, Montpellier et CNRS) qui étudie la manière dont l'élève conceptualise la notion de verbe et l'impact de l'enseignement sur l'évolution de cette conceptualisation, en particulier dans la perspective du passage à l'écrit et la maîtrise d'une langue écrite normée.

Conception, expérimentation et analyse des dispositifs d'initiation à l'écriture de recherche pour la construction de savoirs en didactique de l'écriture

Dans le cadre de la constitution d'un GDR pluridisciplinaire intitulé « Production écrite : apprentissage et expertise » (resp. T. Olive, CR CNRS, CeRCA), un projet sur l'écriture de recherche et la formation de (futurs) professionnels de l'enseignement et de la formation a été élaboré. Ce projet qui s'intitule « Conception, expérimentation et analyse des dispositifs d'initiation à l'écriture de recherche pour la construction de savoirs en didactique de l'écriture » s'intègre dans les axes « Apprentissage & formation, du novice à l'expert » et « Dynamique(s) de la production » du GDR en cours de constitution.

Il s'agira de concevoir, expérimenter et analyser des dispositifs d'accompagnement à l'écriture de recherche d'étudiants en didactique de l'écriture à l'Université (IUFM et autres filières) ou dans d'autres lieux de formation initiale ou continue. Nous appuyant sur les apports de la linguistique, de la psychologie cognitive, de la génétique textuelle et de la didactique de l'écriture, nous considérons l'écriture comme un processus, le rapport à l'écriture (incluant les représentations) comme faisant partie intégrante de la compétence scripturale et le rapport à l'écriture de (futurs) enseignants/formateurs comme déterminant dans l'accompagnement du sujet-écrivain. L'étude se focalisera notamment sur la comparaison des versions définitives et intermédiaires de ces productions afin de repérer les opérations de réécriture (utilisation du logiciel MEDITE, Ganascia, Fenoglio & Lebrave 2004) et d'évaluer leurs effets.

Des collaborations sont prévues avec Bernadette Kervyn (MCF à l'IUFM d'Aquitaine, DAESL, Bordeaux II et chercheur associée à l'ITEM) et Maurice Niwese (docteur en langues et lettres, Université Catholique de Louvain, CRIPEDIS/IACCHOS). La recherche sera articulée à une formation de formateurs (IUFM CVL) sur l'initiation à l'écriture de recherche.

Didactique de l'image

La recherche sur cet axe a pour ambition d'examiner les formes icono-scripturales et les contenus didactiques des discours d'accompagnement pédagogique, à destination des enseignants, conçus selon des modes scripturaux mais censés être communiqués aux élèves selon des modalités pédagogiques orales. La recherche s'inscrit donc dans un continuum écrit de conceptualisation-oral de communication.

La recherche privilégie deux approches complémentaires. La première est sémiologique dans la mesure où les corpus sont décrits comme des ensembles de signes dont le fonctionnement nécessite d'être analysés et les arrière-plans idéologiques dégagés. La seconde est didactique puisque ce sont aussi les choix notionnels qui sont interrogés du point de vue de leur pertinence.

Sont ainsi étudiés des corpus de documents pédagogiques de deux ordres :

- L'un a trait aux situations d'écriture scolaire favorisées par le recours aux images d'albums, en fin de maternelle et en cours préparatoire de l'école primaire.
- Le second est constitué par un corpus de documents écrits à destination des enseignants des années 1950, censés leur permettre de parler aux élèves d'œuvres cinématographiques.

-
La collaboration avec Valérie Vignaux (MCF études cinématographiques, JE 2527 In'Tru (Interactions, Transferts, Ruptures artistiques et culturels) de l'université de Tours, amorcée en 2010 sous la forme d'un projet AAPP Orléans/Tours sera poursuivie. Ce projet donnera notamment lieu à trois journées d'étude.

Formation

Thèses

Projet de thèse (D. Colin, PRAG IUFM CVL) sur le rapport à l'écriture des enseignants (CM2-6^e) (cf. participation au projet déposé par la MSHS de Poitiers).

Master

L'implication du GORDF dans le master Linguistique (spécialité Linguistique et didactique) sera poursuivie. Une implication dans le nouveau master MEEFA (IUFM CVL) est également prévue. Cette double implication ira de pair avec l'encadrement de mémoires en lien avec les projets présentés ci-dessus.

Participation au plan de Formation de formateurs, IUFM CVL

Les résultats des recherches seront également réinvestis dans des actions de formation destinées aux formateurs de l'IUFM CVL.

Le séminaire régulier du GORDF sera poursuivi et maintenu dans ses trois fonctions : confrontation entre chercheurs et diffusion de la recherche, formation initiale SDL et MEEFA, formation de formateurs. Les collaborations nationales et internationales seront privilégiées, en lien avec les projets de recherche du groupe.

2.3 Les programmes de spécialité

2.3.1 Analyse du discours

Responsable : N. Garric

Doctorants : L. Ghanam

Autres participants : J.-Y. Antoine, F. Calas, V. Capdevielle-Mougnibas, M. Goldberg, D. Maurel, H. Maurel-Indart, M. Souchard, A. Tavernier

Les activités de recherche de l'axe Analyse de discours 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 Analyse de discours 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, le groupe Analyse de discours 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.

Depuis plusieurs années, notre équipe de recherche distinguait entre autres deux axes, l'axe Analyse de discours et l'axe Langue et Traitement de l'Information (LTI). Selon les perspectives de recherche propres à chacun, ces axes ont toujours entretenu des collaborations plus ou moins marquées sous différentes formes. Au cours des deux années écoulées de ce quadriennal, elles se sont essentiellement concrétisées par le biais des travaux de recherche des étudiants inscrits en thèse et par le biais d'une coopération scientifique établie entre J-Y. Antoine et N. Garric autour du projet PEPS mené en 2009, du projet DIATEXT ci-dessous défini et de la codirection de la thèse de L. Ghanam. Afin de favoriser ces formes de collaboration, nous avons décidé de ne pas maintenir l'axe LTI afin de l'intégrer partiellement à l'axe Analyse de discours (voir Bilan).

Volet méthodologique

Le volet méthodologique, centré sur la question des données et des procédures qualitatives et quantitatives de traitement de la textualité, maintiendra dans les prochaines années les différentes collaborations entreprises lors du dernier contrat. Deux projets de publication sont actuellement en cours sur les implications méthodologiques de la recherche menée dans le cadre de l'ACI en collaboration avec V. Capdevielle-Mougnibas et l'équipe LISST-CIEU. Une extension de l'analyse est envisagée afin d'élargir le corpus d'étude à des données expérimentales mais également à des données médiatiques relatives aux politiques en matière de formation en alternance, lesquelles sont soutenues par des campagnes de promotion dont les documents seront également étudiés. Cette extension s'appuiera sur le recueil de données orales et écrites et sur la variation des protocoles et des sources. Ce projet est en cours de négociation avec l'Assemblée Permanente des Chambres de Métiers (APCM), notre ambition au-delà d'un financement qui lui serait propre serait d'y associer une thèse CIFRE.

Dans la continuité du projet PEPS mené en collaboration avec H. Maurel-Indart et F. Calas, N. Garric a déposé le projet « DIATEXT : Détection informatisée et analyse textuelle. Les mondes de

l'intertextualité » dans le cadre de l'appel de l'ANR 2010 CREATION associant le LLL à Tours (J-Y Antoine, N. Garric) et Orléans (E. Schang), le CELIS (EA 1002), ICAR (UMR 5191) et le CESR-BVH (UMR 6576).

Le projet DIATEXT est né d'un constat : les évolutions technologiques du numérique favorisent une circulation accrue des œuvres littéraires. Cette mise à disposition est susceptible d'engendrer des difficultés : d'une part, la perte dans le processus de circulation de l'autorité auctoriale propre à un texte, l'appropriation, déclarée ou non, des productions textuelles ; d'autre part, la mise à disposition d'œuvres, qui, livrées au numérique et ainsi décontextualisées, seraient démunies de leur valeur philologique. L'objectif est de transformer les risques inhérents à la circulation numérique des textes en ressources appropriées à la redécouverte du patrimoine littéraire français par une instrumentation logicielle qui identifie, structure et hiérarchise des marqueurs linguistiques pour la caractérisation de la textualité et la formulation d'hypothèses de lecture de la textualité. Le texte est défini comme matérialité complexe dont l'interprétation est construite par des réseaux intertextuels synchroniques et diachroniques multiples tels que des phénomènes de reprise et de détermination hérités, par exemple, d'un genre, d'un courant littéraire ou d'un motif. Le texte, en devenant données, génère une autre forme artistique et de nouvelles modalités de construction de la textualité : il peut être déconstruit et reconstruit à l'aide d'une méthodologie propre jouant des possibilités de recréation des corpus et offrant des variations instrumentales favorables à la définition de marqueurs dont la pertinence naît des mises en corpus elles-mêmes. Cette méthodologie empruntera les ressources du traitement informatique des données textuelles, en particulier celles de la textométrie. La recherche DIATEXT construira une base de données littéraire enrichie de balises XML-TEI à partir de laquelle des corpus seront élaborés pour engendrer une saisie intertextuelle des textes. Des analyses inductives outillées et non outillées, appuyées sur des techniques contrastives, seront menées pour identifier des marqueurs linguistiques de caractérisation textuelle et repérer des faisceaux de marqueurs cohérents pour la formulation d'hypothèses de lecture. En vue de l'instrumentation de la textualité, les marqueurs linguistiques feront l'objet d'une reconceptualisation en adéquation avec les fonctionnalités textométriques. L'instrumentation DIATEXT se concrétisera sous la forme d'un démonstrateur logiciel reposant sur une plateforme Open Source modulable conçue comme une *application* de la plateforme TXM développée par l'ANR Textométrie.

Le volet méthodologique poursuit son investissement pédagogique notamment par un projet de commande en cours d'élaboration en collaboration avec la société INIT, le *Women's forum* et le quotidien *La Tribune*. Ce projet sera consacré à la place des femmes dans le discours patronal, les résultats de l'analyse seront publiés dans le quotidien partenaire. Un projet pédagogique, mené sous la responsabilité d'H. Maurel-Indart et de N. Garric, a également reçu le soutien de l'UFR de Lettres et Langues et des départements de Sciences du langage et de Français. Il permettra l'acquisition de logiciels de traitement textuel textométriques, la formation des enseignants-chercheurs à de nouveaux outils et la mise en place de séminaires destinés au développement des compétences étudiantes à ces outils. Dans le cadre de ce projet, nous avons également prévu de sensibiliser les étudiants de l'UFR de Lettres de Langues à la dimension éthique de la recherche en abordant la question de l'auctorialité scientifique. Nous disposerons dès la prochaine rentrée d'un droit d'utilisation du logiciel de détection du plagiat *Compilatio*. Ce logiciel sera à la disposition des étudiants de master et de thèse afin d'évaluer leur travail de recherche dans le champ de la production scientifique. La recherche des étudiants avec les ressources notamment du numérique repose sur de nouvelles pratiques qui accroissent le risque d'« emprunt » scientifique.

Volet épistémologique

La discursivité, son fonctionnement et ses propriétés, occuperont une place centrale dans le prochain contrat au cours duquel nous interrogerons certains de ses présupposés par l'étude de différents genres discursifs et par la poursuite des investigations entreprises à partir de ce nouveau poste d'observation que fournit le discours de médiation télévisuelle. L'approche n'étant pas typologique dans ce volet, il

s'agit de nous centrer sur des questionnements propres à la discursivité et de les analyser sans objectif de définition générique. Le projet de (re)définir les fondements d'une herméneutique est (re)devenu, dans les analyses de discours mais également dans les sciences humaines et sociales, un objectif scientifique d'actualité. Atteindre cet objectif nécessite, comme l'illustrent de nombreux travaux, une théorisation sémantique mais peut également passer par l'analyse de corpus, tout particulièrement lorsque ceux-ci abordent explicitement le sens comme objet discursif. C'est à cette démarche que nous nous consacrons en étudiant, dans le cadre de l'Analyse de discours de tradition française, comment le discours lui-même peut devenir un lien de théorisation de la discursivité. A cet effet, nous visons l'identification de catégories intradiscursives témoignant de l'activité discursive signifiante des interactants. Celle-ci actualise différentes questions :

- (i) Une fois l'existence de catégories du discours acceptée, comment penser l'articulation de ces catégories aux catégories de langue et comment articuler les deux niveaux ainsi suggérés dans la signifiante ?
- (ii) La question des niveaux d'analyse pose celle des unités, des marqueurs formels du processus de construction du sens. Cette interrogation est ancienne puisque Pêcheux l'avait formulée en restreignant toutefois les formes de la « matérialité discursive » à la syntaxe.
- (iii) Une dernière question se pose, celle de la formalisation du fonctionnement sémantique résultant de l'interaction de ces deux niveaux langagiers et de leurs marqueurs respectifs. Cette théorisation du sens en discours ne peut se satisfaire d'une conception structurale qui en défend l'autonomie et doit s'orienter vers la définition d'une description dynamique du sens inventoriant des parcours interprétatifs pour un signe conçu comme « pluri-accentué » (Bakhtine, Volochinov 1929). La définition d'une sémantique discursive apparaît comme le point nodal de la réflexion.

Notre collaboration avec F. Calas sera renforcée par une collaboration plus large avec plusieurs membres du CELIS de l'université de Clermont Ferrand. Ce projet de recherche adopte pour objectif de cerner les propriétés discursives de certains phénomènes ou de certaines opérations qui mettent en jeu des scénarios de manipulation de la parole. En relation avec F. Calas, nous aborderons cette problématique à partir d'un corpus de discours de fiction mais les travaux menés au sein du LLL l'élargiront à d'autres genres discursifs, le discours médiatique et le discours scientifique notamment. Notre objectif en outre, et la médiation télévisuelle adoptant plusieurs dispositifs dans ses manifestations, sera d'analyser la manipulation tant dans des discours écrits que des discours oraux. Nous souhaitons par cet investissement de l'oral poursuivre les travaux sur le français parlé que nous avons entrepris dans le contrat actuel (Garric & Léglise, 2007).

Volet typologique

Dans le cadre du volet typologique, la collaboration établie entre M. Goldberg et N. Garric sera maintenue au sein d'une recherche qui associera les travaux sur la scientificité à ceux consacrés à la médiation au sein d'un projet à l'examen, soumis à l'appel PEPS-2010, « Les débats de société à thème scientifique » porté par N. Garric et associant M. Goldberg, A. Tavernier et M. Souchard. Ce projet s'appuie sur le développement de pratiques discursives d'association des citoyens aux communautés expertes, scientifiques, politiques, économiques, notamment illustrées par les conférences citoyennes, les débats publics, la médiation médiatique ou encore les forums de discussion pour interroger le fonctionnement de la discursivité. L'émergence de cette parole démocratique est susceptible de définir un nouvel espace de construction du savoir et des connaissances. Ce projet souhaite en interroger les pratiques en sélectionnant comme objet exploratoire des débats de société à thème scientifique. Nous développerons l'analyse par un certain nombre de questions très précises qui, visant les enjeux d'un genre, portent sur des tensions construites par le dispositif discursif lui-même : la scientificité *versus* la vulgarisation ; l'objectivité concernant l'objet scientifique *versus* la subjectivité du scientifique ; l'impartialité *versus* l'engagement ; l'information *versus* le commentaire et le disciplinaire *versus* l'interdisciplinaire.

2.3.2 Programme de Recherche sur *Oralité, Histoire, Ecriture* dans le Monde Ibérique, d'Orléans / PROHEMIO

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

Doctorants : C. Hounnouvi, Y. Haquin

Collaborations : J. Fernandez, M. Feugain, F. Giraud

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'inscrivent toujours sur ces deux axes et en cela ils s'intègrent parfaitement à l'orientation actuelle du Laboratoire Ligérien de Linguistique qui met l'accent sur l'étude de l'oral. L'objectif du PROHEMIO est donc d'étudier cette problématique de l'oralité portée par le laboratoire en l'appliquant au domaine ibérique et en interrogeant des approches de l'oralité originales qui s'articulent autour de deux grands axes.

Oralité et écriture : étude du figement linguistique

Le PROHEMIO étudie les rapports entre oralité et écriture sous l'angle du figement linguistique en raison du statut très particulier de ces expressions naviguant entre oral et écrit non seulement par leur emploi (qui fait la part belle à l'oral) mais aussi par leur genèse. Le PROHEMIO s'intéresse au figement sous trois angles différents :

Le figement : analyse linguistique et traduction.

Poursuite du *Dictionnaire idéologique comparatif des locutions espagnoles et françaises* sous la direction de Sylvie Fournié-Chaboche. Le travail de réflexion mené jusqu'en 2010 a permis d'élaborer un projet neuf en introduisant une dimension idéologique (on commencera par un dictionnaire des défauts) et en envisageant un support informatique (réalisation du dictionnaire à l'aide du logiciel Toolbox). Cela permettra de publier plus facilement et fréquemment des parties du dictionnaire, à mesure de sa constitution, en l'hébergeant sur le site de l'Université ou tout autre site. Le travail envisagé durant ces quatre prochaines années se décomposera de la façon suivante :

- Poursuite du travail de constitution d'un corpus sur des textes espagnols non traduits et traduits en français (cf. recension en annexe). Il est prévu d'étendre les investigations à la presse et à la publicité, (thèse de Christian Hounnouvi).
- Application du logiciel Toolbox pour une mise en parallèle du trésor lexicographique des deux langues. Au plan méthodologique, le dictionnaire sera classé de façon thématique. A l'intérieur d'un thème, on distinguera les principales nuances liées aux différentes possibilités de sens littéral. De même, les différentes locutions seront elles-mêmes classées par type d'images auxquelles elles renvoient. De part et d'autre de cette combinaison sens/image (qui constituera l'ossature du dictionnaire aux dépens de la traditionnelle classification alphabétique) figureront ou non les locutions présentes ou non dans chacune des deux langues.
- Rédaction du Dictionnaire (en lien avec la thèse de Yohan Haquin).

Le PROHEMIO envisage de réaliser, dans un premier temps un dictionnaire des défauts, cette thématique ayant semblé très féconde au niveau phraséologique. Au-delà de l'intérêt purement traductologique, la mise en perspective des deux langues dans le domaine des défauts sera certainement propice à des analyses sur les différentes représentations et modes de pensée de ces deux sociétés a priori très proches et ce, par le prisme de l'oralité. Celles-ci pourront donner lieu à publications à mesure de l'élaboration du dictionnaire.

le figement linguistique dans le discours littéraire

Le PROHEMIO se consacre aux relations entre oralité et discours littéraire avec la publication de nombreux articles, l'organisation de plusieurs colloques, la soutenance en 2002 de la thèse de Sylvie Fournié prolongée par la publication de différents articles. Le PROHEMIO développera au cours du nouveau contrat un questionnement relatif à la traduction des formes fixes en discours, notamment dans le discours littéraire.

Naming et analyse sémiotique du figement linguistique dans la communication d'entreprise

Dans le cadre de l'articulation entre la formation (notamment dans des filières professionnelles) et la recherche, le PROHEMIO a commencé à développer cette approche innovante du figement à travers l'analyse du *naming*, des slogans, et de manière générale de la communication d'entreprise. Cet axe de recherche, illustré par des publications (Fournié-Chaboche) et la préparation d'une thèse (Ch. Hounnouvi) recèle un fort potentiel d'application au regard du dynamisme de ce secteur d'activité économique.

Oralité, histoire et culture traditionnelle

Recueil de témoignage oraux

Les projets de recherche pour les années à venir se centrent principalement sur la poursuite des travaux en lien avec la Memoria Histórica. L'accès récent à des documents concernant la région de Ciudad Rodrigo (Salamanque), et dont nous ignorions l'existence, nous a conduit à élargir le secteur géographique de nos investigations et nous sommes pressés de recueillir les témoignages oraux des personnes qui ont vécu cette période de la Guerre Civile espagnole, car elles sont de moins en moins nombreuses et fragiles. Une première publication sur ce sujet devrait voir le jour prochainement suivies d'autres au cours du prochain contrat.

Organisation de colloques

Une autre facette de l'activité du PROHEMIO consiste en l'organisation de colloques et en la publication d'actes. Après avoir consacré beaucoup de colloques au figement, le PROHEMIO aborde à présent dans ces colloques la question de l'oralité en lien avec l'histoire (guerre civile espagnole) et la culture populaire. Les colloques du PROHEMIO ont lieu tous les deux ans (les années impaires). La publication dans les *Cahiers du PROHEMIO* (Université d'Orléans) a lieu l'année suivante. Le neuvième colloque du PROHEMIO sera organisé en Espagne en juillet 2011. Le thème fixé est « Miscelanea rebollana », toujours dans la perspective du PROHEMIO depuis sa création, c'est-à-dire en étudiant les rapports entre oralité et écriture, suivant les trois axes, histoire, langue et culture traditionnelle.

Diffusion de la connaissance: transmission d'un patrimoine oral

Après la très bonne réception des conférences données cet été dans plusieurs localités du Rebollar : Robleda (10 et 11 août 2010), El Bodón (17 août 2010) et Villasrubias (22 août 2010), à l'occasion du 1100 anniversaire de la création du Royaume de León, sur le thème « Lengua e historia : 1) El reino de León, las hablas leonesas y El Rebollar ; 2) El habla de El Rebollar y las hablas vecinas de Extremadura », il est envisagé la mise en place de cours sur *el habla de El Rebollar*, pour l'été 2011, cours destinés à un public étudiant mais ouverts à toute personne intéressée, dans une perspective inter-générationnelle, avec la participation de personnes âgées parlant encore actuellement le dialecte étudié.

2.4. Les thématiques transversales

2.4.1 Perception et ordonnancement

Responsable : S. Osu

Au cours des quatre prochaines années, les investigations de la thématique transversale, conduites sur les thèmes de la dénomination et de la temporalité, seront prolongées par un travail collectif sur l'ordonnancement des unités linguistiques afin de cerner, d'une part, le rôle non seulement du temps mais aussi du sujet (parlant/interprétant) dans l'agencement des unités ; et d'autre part, les principes, si principes il y a, qui actualisent préférentiellement tel ou tel mode d'ordonnancement, aussi bien dans chacune des langues représentées qu'à travers les langues, dans leur diversité.

Réunissant l'ensemble des chercheurs du LLL sur le site de Tours, EC, EC associés, ATER, doctorants voire étudiants deM2, ce programme poursuivra une réflexion sur la thématique à partir de l'étude des langues représentées dans les compétences du laboratoire (français, anglais (et vieil-anglais), russe, allemand, dagara, ikwéré) et des langues des collègues invités à s'associer au travail mené lors de journées d'études et de séminaire. La question de l'ordonnancement interpelle directement les informaticiens qui contribuent à l'animation de l'équipe.

L'étude concernera plus particulièrement :

- l'ordre des mots (déterminant/déterminé ou déterminé/déterminant),
- l'ordonnancement des discours (publicitaires, médiatiques, politiques, etc.),
- le rôle spécifique des accents dans l'agencement des unités,
- la perception qu'un sujet (parlant et/ou interprétant) peut avoir d'un ordre des unités donné,
- et à partir des travaux de la linguistique historique et diachronique, l'ordre suivant lequel les unités d'une langue donnée paraissent évoluer.

Comme cela a été le cas pour les précédentes thématiques, l'organisation d'un colloque ou de journées d'études est envisagée même si, eu égard à la récence des études déjà réalisées, il est difficile de donner une date dans le cours du prochain quadriennal (2012-2015).

2.4.2 Sémantique instructionnelle et prosodie

Participants LLL : P. Cadiot, G. Col, F. Nemo

Participants LATTICE : J. Aptekman, T. Poibeau, B. Victorri

Doctorants et Post-Doc : S. Girault, E. Grimaud, M. Petit

Le projet SIP (Sémantique instructionnelle et prosodie) est développé en commun par le LLL et Lattice ([Langues, Textes, Traitements informatiques et Cognition, UMR 8094 du CNRS, ENS Paris – voir annexe). Il vise à réaliser une modélisation du sens des marqueurs discursifs de l'oral prenant en compte l'intégralité du contexte de leurs emplois : non seulement les constructions syntaxiques et macro-syntaxiques et la sémantique des unités lexicales co-textuelles, mais aussi la prosodie, qui joue un rôle essentiel dans la détermination du sens précis des marqueurs prosodiques.

Objectifs

Au plan théorique, il s'agit d'approfondir la connaissance du comportement sémantique et du rôle des marqueurs dans le discours, et de mieux comprendre l'interaction de la prosodie avec les autres niveaux linguistiques de construction du sens.

Au plan applicatif, l'objectif est de mettre au point des outils permettant de calculer le sens précis des marqueurs en contexte, ces outils pouvant jouer un rôle essentiel dans des tâches de traitement automatique (extraction d'information, systèmes de questions-réponses, etc.). Pour ce faire, à chaque marqueur est associée une instruction unique qui régit son comportement sémantique en contexte. Cette instruction comporte des variables qui doivent être obligatoirement instanciées par le contexte (au sens large : depuis le co-texte de l'énoncé jusqu'aux données de la situation d'énonciation et des connaissances partagées des interlocuteurs). Les gestes prosodiques sont décrits par des instructions au même titre que les autres éléments linguistiques constituant les énoncés.

Méthode et plan de travail

Pour chaque marqueur étudié, une banque de données de ses emplois dans des corpus oraux de parole attestée est constituée, combinant plusieurs genres (parole spontanée, interviews, discours, etc.). Chaque emploi est caractérisé par des annotations morpho-syntaxiques, sémantiques, discursives et prosodiques. La modélisation proposée pour un marqueur est systématiquement confrontée à sa banque d'emplois, qui constituera l'outil de validation des hypothèses.

2010 : livraison d'une première version d'un logiciel d'annotations (Analec) et mise en place d'une banque d'emplois pour les marqueurs *alors* (en français) et *well* (en anglais).

2011 : confrontation des modélisations proposées pour ces deux marqueurs et mise en place d'un travail d'annotation de grande envergure sur une dizaine de marqueurs (en français et en anglais).

2012 *sq.* : extension à de nouveaux marqueurs et de plus grands corpus.

2.4.3 *Sémantique des langues, points de vue et gestion des connaissances*

Responsable : P.-Y. Raccach

Ce programme interdisciplinaire traite de questions relevant principalement des sciences du langage, des sciences de la cognition et des sciences de l'information ; il s'appuie sur les résultats des travaux de recherche de ces trois dernières décennies.

Une ambition théorique.

- (i) *Du point de vue linguistique* : décrire rigoureusement et empiriquement les phénomènes relevant de la contribution que les langues apportent à la construction du sens, en se fondant sur l'observation des effets des énoncés ; parmi les principaux éléments qui seront pris en compte pour étudier cette contribution, figurent : le lexique, les articulateurs, la prosodie. Cette démarche, fondée sur l'idée que les unités des langues imposent des contraintes sur les points de vue que leurs énoncés peuvent évoquer, permet d'utiliser les structures sémantiques comme modèles de structuration des connaissances pour la simulation. La description sémantique des unités prosodiques permet de ne pas limiter cette démarche au traitement de l'écrit.
- (ii) *Du point de vue des sciences de la cognition* : l'examen approfondi de l'hypothèse d'une analogie de structure entre la gestion humaine des connaissances et celle du sens, analogie qui s'observe dans l'étude des instructions que la langue fournit pour construire le sens des énoncés, constitue une alternative importante à la réduction du concept de connaissance à celui d'informations..
- (iii) *En intelligence artificielle* : élaborer et mettre en œuvre des concepts permettant de simuler l'expertise dans son fonctionnement, et pas seulement en utilisant la puissance de calcul pour obtenir des résultats proches de ceux que des experts obtiendraient.

Une ambition pratique

- a) *En linguistique* : construire un dictionnaire plurilingue des idéologies inscrites dans les lexiques. Ces idéologies, que les mots marquent, apparaissent comme des points de vue imposés par les discours, et se manifestent par des biais sociocognitifs que tout observateur de ces discours peut repérer.
- b) *En sciences de la cognition* : construire des outils conceptuels permettant la représentation des connaissances expertes et leur recueil à partir d'énoncés et de discours d'experts. Ces outils, c'est l'objectif du programme, seront fondés sur un modèle commun devant être adapté à la fois à la représentation sémantique et à la représentation des connaissances expertes, et seront utilisés (et donc testés et affinés) en entreprise, dans le cadre d'actions de valorisation.
- c) *En intelligence artificielle* : s'appuyer sur les concepts élaborés dans le cadre de la recherche théorique pour construire des outils informatiques d'aide au recueil et à la gestion des connaissances, notamment, dans le cadre d'une approche de simulation de l'expertise, mais aussi dans le cadre d'approches de services aux entreprises.

Cette approche est confortée par l'absence de distinction opérée entre les points de vue relevant de connaissances expertes et ceux qui relèvent d'identifications idéologiques au niveau de l'analyse linguistique des discours. La discrimination entre les deux types de points de vue nécessite de recourir à des informations extralinguistiques (statut des inférences, statut cognitif des interlocuteurs, statut social, objectifs...). Cela suggère un homomorphisme entre les structures de représentation des points de vue, quelle que soit leur nature, ce qui a conduit à l'élaboration de la *sémantique des Points de Vue*.

2.4.4 Phonologie

Responsable : G. Bergounioux

Participants : V. Abasq, J.-M. Fournier, J.-L. Rougé, E. Schang

Doctorants : X. Luo, M. Martin

Autres participants : A. Delplanque, X. Yu

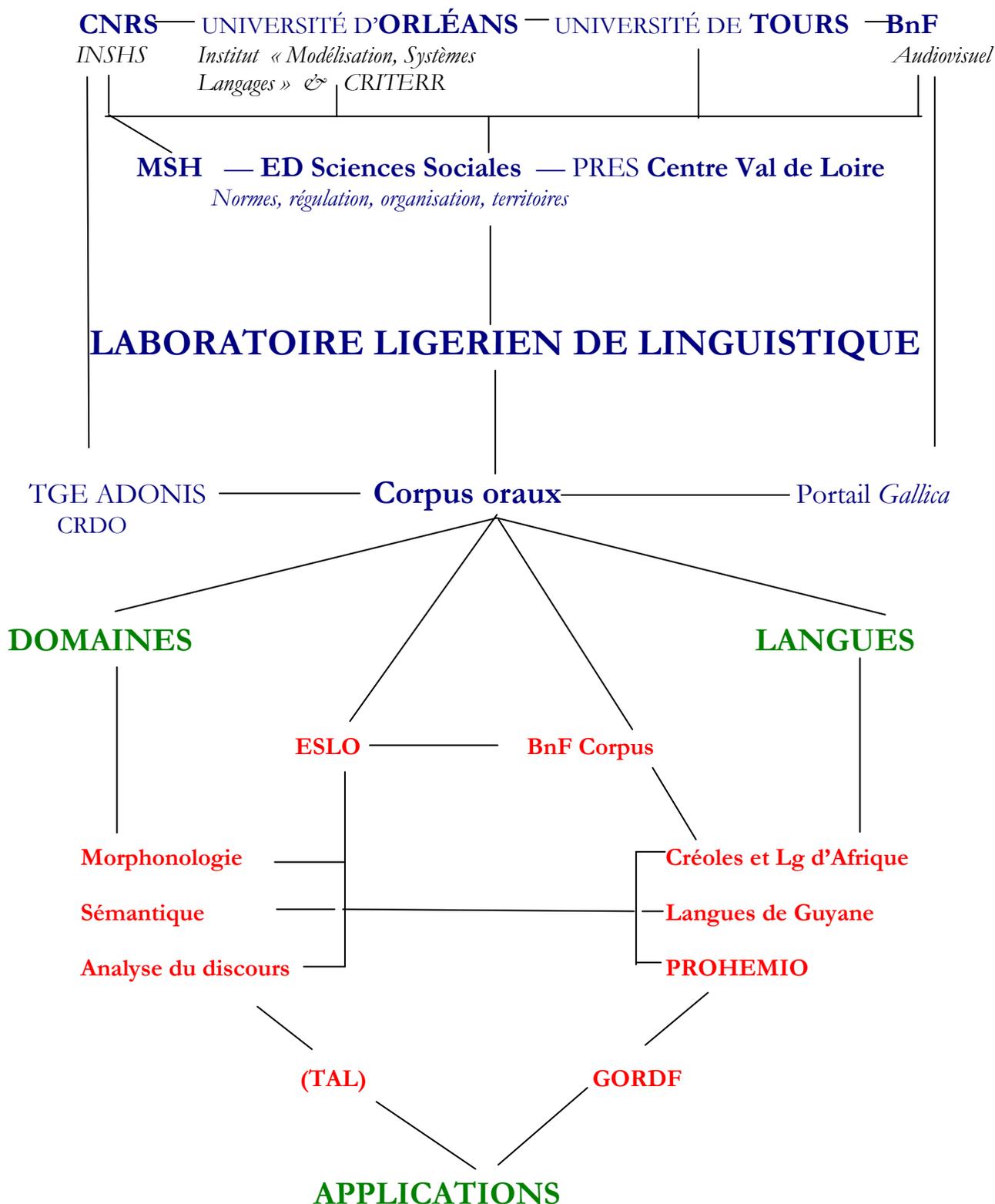
Ce programme articule les études phonologiques à Tours (en particulier les travaux conduits par l'équipe morpho-phonologie) et à Orléans, qu'elles aient pour objet la notation des langues en linguistique de terrain ou la formalisation linguistique.

Ces études se proposent trois objectifs :

- une réflexion théorique qui a notamment conduit à orienter le colloque annuel du CERLICO, pour les deux années 2009 et 2010 sur un thème central de la description des langues : « Transcrire - Ecrire, - Formaliser ». La réflexion se poursuivra (i) autour de la question de la transcription restituée de la forme orale des langues impliquée par les projets de corpus, (ii) d'autre part autour de la notation des langues peu ou pas décrites, en Amérique et en Afrique (en particulier au Burkina), voire autour des variétés émergentes par créolisation ou formes instables du français urbain populaire dans les grandes villes de l'Afrique subsaharienne (thèse de M. Cavalheiro), enfin (iii) par des propositions de notations des langues (y compris dans les outils métalinguistiques) ;
- une application et une extension aux problèmes que posent les notations et les transcriptions aux applications informatiques et à la mise en ligne, à l'exploitation par le traitement automatique des langues,
- une activité d'animation scientifique dont l'exemple vient d'être donné avec la relance, entre Orléans et Tours, du Réseau Français de Phonologie (voir annexes). Héritier du GDR « Phonologies » animé par B. Laks, le RFP était tombé en sommeil depuis 2005 au grand regret de la communauté des chercheurs, en particulier des doctorants, qui y avaient trouvé le cadre idéal d'exposition de leurs recherches. Avec un comité scientifique renouvelé et élargi, le RFP a repris ses travaux avec un colloque en juillet 2010 à Orléans qui sera suivi d'un second à Tours en 2011, d'un troisième à Paris en 2012. Un site électronique a été ouvert à l'initiative de Roland Noske (Lille 3) et d'autres initiatives doivent être discutées à l'AG prochaine du Réseau.

Ce programme, transversal par définition, est largement impliqué dans les formations doctorales, à Tours (M. Martin...) comme à Orléans (X. Luo). Il vient en appui disciplinaire dans les formations de licence et de Master, de sciences du langage et de langues.

ORGANIGRAMME DE LA RECHERCHE



3. Prospective

La suite des discussions avec le département de l'audiovisuel à la BnF, l'impulsion apportée par l'ILF et les demandes du CNRS, en particulier du TGE ADONIS, les collaborations internationales également, détermineront les travaux conduits au cours du prochain contrat et dans la décennie à venir. L'accélération des opérations de numérisation, la demande exponentielle de contenus, la gestion des bibliothèques virtuelles et des bases de données, les opérations d'enrichissement, l'accessibilité des chercheurs aux matériaux et du grand public aux sources orales de sa mémoire situent la démarche entreprise au centre d'un redéploiement des techniques et des méthodes d'analyse.

Le LLL n'a pas les moyens de répondre à un enjeu d'une telle importance ; il a les capacités d'inventer, avec l'expertise des conservateurs et des chercheurs de la BnF, sur des segments particuliers et sur des corpus collectés et traités de première main qui serviront de banc d'essai, des procédures et des méthodes de travail qui prennent en compte l'ensemble du process. L'objectif est, dans le cas du français et des langues représentées au sein du laboratoire, de contribuer à la production du grand corpus numérique du français, mais aussi de langues d'Afrique et de langues d'Asie, créoles et langues mixtes, dont des éléments ont été collectés et exploités à Aix-en-Provence, Lyon, Montpellier, Nice, Paris 3, Paris 10, Paris 11, Perpignan, Poitiers, Saint-Denis, Toulouse et bien sûr en Belgique (Louvain), au Québec (Montréal, Sherbrooke). Cette liste n'est pas exhaustive. Elle prouve qu'un champ de recherche très actif s'est constitué dans une démarche ascendante qui a privilégié la collecte et l'exploitation scientifique sur la conservation patrimoniale et la diffusion. Des questions comme la pérennité des supports, l'interopérabilité, la communauté des annotations requièrent un effort collectif pour lequel la BnF, le CRDO et le TGE ADONIS sont des interlocuteurs privilégiés. Avec les fédérations (ILF et TUL), la Délégation Générale à la Langue Française et aux Langues de France (DGLFLF / Observatoire des Pratiques Linguistiques), l'Agence Universitaire de la Francophonie, les Archives nationales, des ressources devront être mis en réseau, qui associeront les universités et le CNRS. Le LLL s'inscrit pleinement dans ce projet où il entend, à l'échelle inter-régionale du centre et de l'ouest, jouer un rôle moteur.

Les prochaines années doivent permettre de capitaliser les recherches conduites lors du précédent quadriennal, non seulement dans le programme ESLO mais aussi dans l'exploitation de bases de données dictionnaires (morphophonologie), dans la phase de documentation des langues (langues d'Afrique et de Guyane), dans le champ de la sémantique et de la prosodie, de la sociolinguistique et de l'épistémologie... Intégré à la MSH du Centre Val de Loire et en synergie avec les chercheurs de la BnF, le LLL attend de ses tutelles un accompagnement en postes de chercheurs, d'ingénieurs et de personnel administratif et en financement.

Fort de cet appui, des perspectives encourageantes se dessinent pour créer un centre régional d'expertise en matière de conservation patrimoniale en partenariat avec les laboratoires du CNRS. Parallèlement, un renforcement des collaborations avec les laboratoires d'informatique, déjà très active avec le L.I. de Tours et qui monte en puissance à Orléans comme le montre l'organisation de l'institut « Modélisation, Systèmes, Langages », devrait permettre des transferts dans le domaine technologique, que ce soit dans la mise à disposition d'une compétence à des nations émergentes qui se trouvent confrontées de façon aiguë à la pérennité de la mémoire ou vers les industries de la langue.

La maîtrise de ce champ par la recherche publique est aussi une réponse au défi de l'appropriation des corpus oraux. Après l'opération de numérisation des corpus écrits par Google, qui ont appelé une réponse des bibliothèques européennes, la réponse pour 2020 est à présent du côté de l'oral.

4. Une organisation administrative et financière à définir

4.1 Administration

L'obtention du statut d'UMR conditionne l'ensemble des autres points :

- définition des statuts d'appartenance (en particulier pour les personnels de la BnF),
- composition du corps décisionnel de l'AG et des assistants,
- mode d'élection du conseil de laboratoire.

Ces dispositions devront correspondre aux attentes de chacun des partenaires dans le respect de ses spécificités, thématiques et géographiques,

Il est prévu la rédaction d'un règlement intérieur, discuté entre l'ensemble des partenaires et ratifié en AG avec validation par les tutelles (CNRS, universités et BnF). Le conseil de laboratoire remplacerait le bureau et sa composition sera définie en sorte qu'y soient représentés les personnels CNRS et les conservateurs et les autres personnels de la BnF, ainsi que les personnels techniques travaillant pour le laboratoire attendus au cours du prochain contrat.

Pour son fonctionnement, le LLL reprendra les statuts correspondant aux usages des UMR, en conformité avec les demandes du CNRS et du Ministère de la Culture et de la Communication, en veillant à la représentation des personnels techniques et des doctorants. Le processus décisionnel de l'AG et du conseil de laboratoire seront précisés à ce moment. La différence des statuts et des missions ne permet pas de décider dès à présent l'ensemble des modalités d'affiliation en sorte que figure, dans ce dossier, et sur le même plan, l'ensemble des enseignants-chercheurs ayant demandé leur affiliation au laboratoire. Figurent comme associés ceux qui ont un autre rattachement principal. Aucun collègue d'une université étrangère n'a pour le moment été compris dans cette catégorie afin de ne pas alourdir la nomenclature. Il en est de même pour les ATER et post-doc qui figurent simplement à titre de collaborateurs.

4.2 Financements

Les financements sont attendus de cinq bailleurs :

- les tutelles (BnF, CNRS, universités de Tours et d'Orléans)
- les programmes de recherche (ANR, ERC...),
- les collectivités publiques,
- les appels d'offre hors agence,
- les entreprises.

Concernant les tutelles, et en fonction de la politique de chacun des partenaires, un abondement annuel doit demeurer la base prévisionnelle pour la conduite d'une politique suivie. Dans les universités, un versement correspondant à 2.000 € par EC et à 1.000 € par doctorant est aujourd'hui considéré comme un plancher pour un fonctionnement qui peut moduler cette somme autour d'un critère qualitatif. Les dotations CNRS viendront comme un complément indispensable aux ambitions nouvelles du laboratoire. La part de la BnF, au-delà des engagements pour ses personnels, est comprise dans la négociation globale avec cet établissement et avec la direction de la recherche au MCC.

Une attente forte devra trouver sa concrétisation au prochain contrat : la nécessité d'affecter les moyens en personnel administratif et informatique au laboratoire afin de permettre un fonctionnement normal de l'unité. Il n'est pas de bonne politique de soustraire du temps aux EC afin qu'ils se consacrent à des travaux de secrétariat, de comptabilité et de gestion pour lesquels ils sont peu compétents et encore moins disponibles. La carence en informatique aura été l'un des plus graves

obstacles auxquels se soit heurté le LLL durant l'actuel contrat quadriennal contraignant, à Tours comme à Orléans, à des prestations de service externes coûteuses (elles ont représenté l'équivalent d'une année et demie de l'abondement récurrent par les deux établissements) L'absence d'IE dans l'équipe représente un appauvrissement comparatif. De même, il est regrettable qu'un certain nombre de tâches d'administration aient dû être financées par des CDD occupés par des étudiants en Master ou en thèse dont ce n'est pas normalement la vocation.

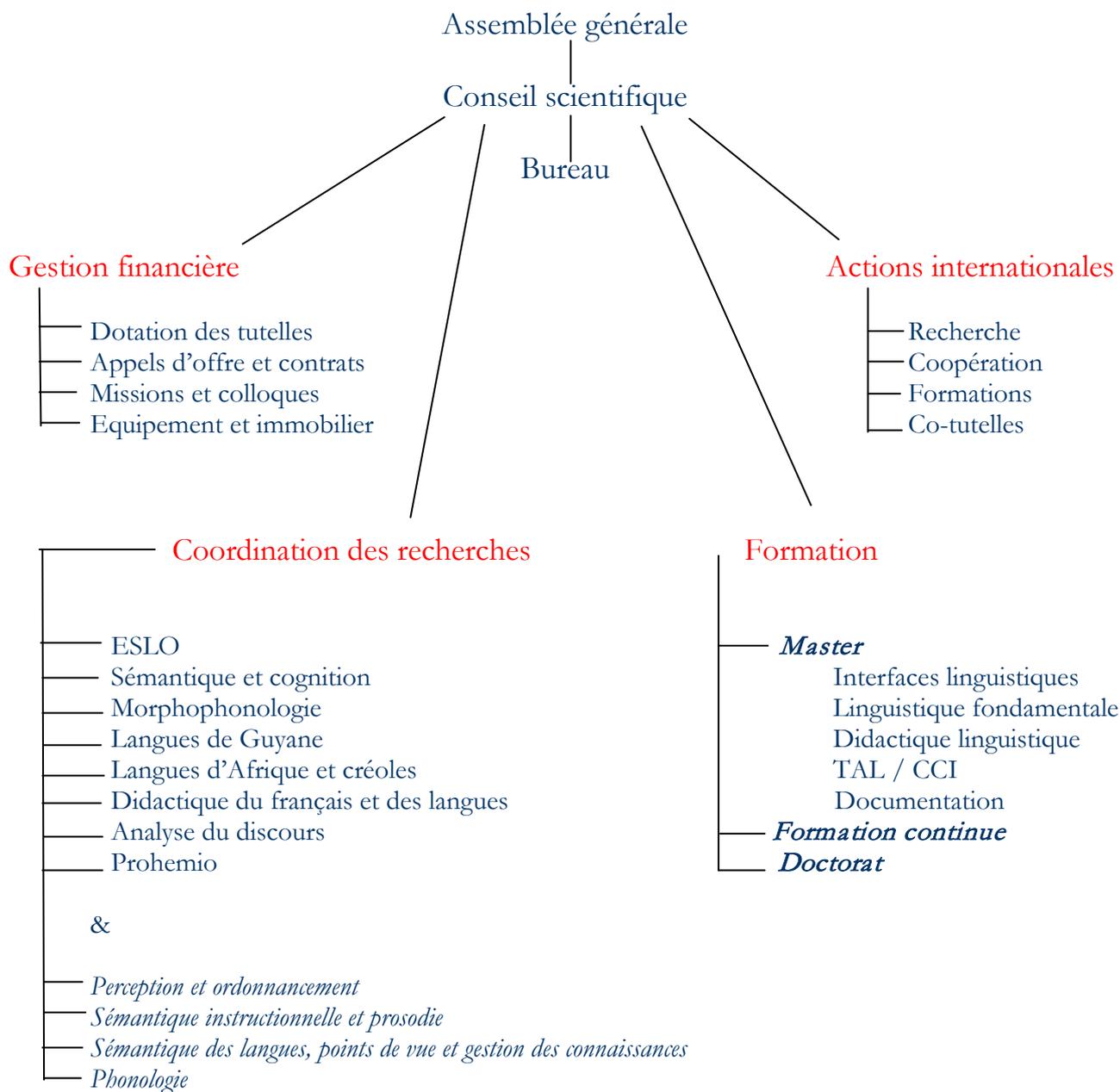
Pour les appels d'offre, qui ont représenté plus des deux tiers des lignes de crédits disponibles dans le laboratoire, qu'ils émanent des universités, de la région Centre, des agences ou des Ministères, ils seront sollicités comme ils l'ont été même si la lourdeur d'un certain nombre de dossiers s'avère parfois dissuasif, en particulier dans le cas de l'ERC. Le LLL, qui a été sélectionné dans l'appel Corpus 2006, entend s'inscrire dans de prochains appels en trouvant les partenariats les mieux adaptés à sa nouvelle thématique.

Les collectivités publiques ont apporté un soutien important en finançant des contrats doctoraux (6,5 par an pour les deux établissements), des programmes de recherche (CRITERR), des appels d'offre (LCO, PANGLOSS) et en subventionnant différentes manifestations scientifiques, en particulier les colloques. Pour un laboratoire qui intervient à une échelle régionale, il s'agit d'un partenaire indispensable avec lequel sera poursuivi un dialogue fructueux et parfois complexe.

Concernant les ressources sur appels d'offre hors établissement et hors agences, elles sont venues du CNRS (ILF, programme Oyapock, PEPS), de l'AUF et de la DGLFLF/MCCC. La thématique des corpus oraux et le projet d'un corpus de référence du français devrait permettre de soumettre des propositions qui correspondent aux missions et aux attentes de ces différentes instances et, pour la partie pédagogique, du Ministère de l'Education.

Enfin, pour des applications, un soutien d'entreprises innovantes, en particulier dans le secteur des industries de la langue et du didacticiel, est attendu même si l'étroitesse du marché, la faiblesse de ce secteur dans le Centre et la difficulté de mobiliser sur des thématiques qui ne soient pas seulement des applications ont jusqu'à présent largement réduit ce que pourraient être les retombées du CIR dans ce domaine.

ORGANIGRAMME ADMINISTRATIF



Le nom des directeurs de programme figurent dans ce dossier. L'implication des collègues de la BnF sera complétée suivant leurs indications et les décisions prises par l'assemblée générale du laboratoire.

5. Indicateur des objectifs et échéancier (2012-2015)

5.1 Structuration du laboratoire : intégration et encadrement

Dès que les négociations avec la BnF l'auront permis, une nouvelle phase de discussion sera engagée avec le CNRS, impliquant l'ILF et le TGE ADONIS, afin de convenir avec l'ensemble des partenaires et les deux universités le type de structuration à envisager. Il est délicat d'anticiper cette étape qui préjugerait du résultat d'une négociation qui, si bien engagée qu'elle se présente, n'est cependant pas au terme d'un parcours administratif qu'on sait complexe par nature et riche en rebondissements.

Les compétences combinées des deux partenaires, le LLL dans un travail sur des corpus oraux à partir de langues différentes, au premier plan le français (quatorze chercheurs) et l'anglais (8 chercheurs), et dans des disciplines complémentaires, la BnF dans le traitement des documents sonores demande à être articulé dans un travail concret qui a conduit à mettre en avant la notion de process comme enchaînement d'opérations successives incrémentant l'information sur les données exploitées. L'articulation administrative et fonctionnelle entre deux administrations intégrées à des ministères distincts et dont les missions sont souvent différentes supposera un temps d'adaptation et une définition des méthodes de travail que la distance complique. L'expérience d'une présence sur deux sites distincts de cent kilomètres a donné à ce sujet une première expérience qui a démontré qu'il n'y avait rien d'insurmontable dans un tel projet.

Un prolongement de ce travail est attendu d'une politique de formation qui, reprenant ce qui existe dans les universités, autant dans les masters (y compris un master de documentation en discussion avec l'IUFM) qu'en doctorat, et à la BnF, veillera à enrichir les contenus d'enseignement par les travaux conduits en collaboration. Cette politique vaudra non seulement pour les étudiants en thèse, voire en Master, associés aux programmes du laboratoire mais aussi pour les personnels de la BnF qui souhaiteraient bénéficier de VAE, s'inscrire en recherche, et pour les MCF qui entendraient passer une HDR, un objectif indispensable à réaliser dans le LLL où plusieurs chercheurs publiants envisagent de se qualifier aux fonctions de professeur.

5.2 Partenariats : réseau interrégional, réseau thématique, réseau international

Au-delà de l'intégration sur chacun des sites et des activités dans le PRES Centre Val de Loire, le LLL entend être un acteur dans la structuration régionale des SDL, dans les organisations thématiques, jusqu'au niveau national et international.

Une collaboration privilégiée avec l'Université de Poitiers a été engagée à travers différents programmes concernant les corpus oraux, la didactique et les études sur l'oral. Impliquant en particulier la MSHS et les échanges avec le FORELL, ces coopérations réunissent l'essentiel du potentiel des deux régions Centre et Charentes-Poitou. Au-delà, il semble que ce soit le CERLICO, où le LLL est doublement représenté, qui soit la structure la plus adaptée à une projection inter-régionale (Limousin, Bretagne, Pays de Loire et Basse-Normandie) où la présence d'une UMR aurait un rôle moteur.

Sur un plan thématique, ce sont les programmes de recherche qui ont décidé de certaines collaborations privilégiées, en particulier avec le LATTICE et, ponctuellement, avec Modyco (Paris X), le LIDILEM (Grenoble), LLF (Paris 8). Le LLL s'est inséré de plein droit dans les études conduites en sémantique à travers différents programmes et publications (sémantique linguistique, sémantique cognitive...), en syntaxe en intégrant le programme de recherche sur les grammaires créoles et en phonologie en relançant le Réseau Français de phonologie. L'intervention sur des aires linguistiques communes a conduit à des discussions avec le LLACAN.

Au niveau international, les collaborations suivies avec des chercheurs d'une dizaine de pays

(Angleterre, Belgique, Brésil, Burkina, Canada, Cap-Vert, Chine, Espagne, Portugal, Roumanie, Tunisie), de façon plus conjecturale avec ceux d'une demi-douzaine d'autres (Allemagne, Australie, Etats-Unis, Haïti, Maroc, Pologne, Sénégal, Suisse) se heurte à la différence des structures et des modes de fonctionnement.

Afin de ne pas disperser entre des accords nombreux et peu substantiels ses efforts, le laboratoire se concentrera sur des relations ciblées autour de quelques partenaires et sur des thématiques fortes : les corpus oraux en Europe, la description des langues au Burkina, la formation en Tunisie, la didactique du français avec le Canada et la Belgique par exemple.

5.3 Investissements et financements

En s'investissant dans la constitution, l'archivage, la conservation, l'étude et la mise à disposition des corpus oraux, le LLL s'engage dans un domaine où les coûts sont sans commune mesure avec ceux mis à la disposition d'un laboratoire de SHS. C'est l'une des raisons pour lesquelles deux types de collaboration ont été prévus.

Concernant l'archivage, la maintenance et, pour une part, la diffusion, l'expérience irremplaçable du Département de l'Audiovisuel de la BnF offre des ressources sans équivalent en France. Concernant la préservation, le stockage et l'exploitation en recherche, les discussions engagées avec le TGE ADONIS doivent permettre de trouver des solutions à l'usage des corpus.

D'autres synergies sont envisagées sur le plan régional (avec la MSH, les laboratoires liés à la conservation – CESR, CITERES, IRAMAT, IRHT), national, avec les Fédérations, en particulier l'ILF, et les UMR impliquées dans une démarche semblable, au premier rang desquelles, pour son rôle dans le CRDO, le LLP (Aix-en-Provence), et pour son expertise dans l'exploitation des corpus ICAR (Lyon), de même, dans les corpus oraux, l'ATILF (Nancy). Enfin, au plan international, les expériences de plusieurs pays, en particulier le Royaume-Uni, paraissent particulièrement importantes.

Le recours à différentes infrastructures devrait permettre de contenir le budget informatique dans des proportions qui, pour importantes qu'elles soient, bénéficiera de mutualisations. Il n'en reste pas moins qu'il y aura un coût, en matériel et en logiciels, qui devra être pris en compte, en particulier pour assurer l'interopérabilité des serveurs et assurer la sécurité et la maintenance.

Echéancier

2012

- Intégration dans le TGE ADONIS
- Mise en place d'une plate-forme informatique
 - interaction Universités de Tours et d'Orléans / BnF
 - intégration serveur CNRS
- Mise en place d'une structure d'édition électronique

2013

- Développement de la plate-forme informatique
 - participation aux initiatives BnF sur les corpus oraux : journées d'étude...
 - interconnexion avec les laboratoires (via l'ILF ?)
 - construction d'un serveur pour travail à distance avec le Burkina

2014 & 2015

- Contribution à la partie « documents oraux » de la grande bibliothèque numérique
- Préparation d'une interconnexion au niveau européen

En terme immobilier, le principal investissement concerne le site de Tours qui ne dispose pas des locaux correspondant aux ambitions de l'équipe. Les autres dépenses sont celles d'un laboratoire intervenant dans le domaine des SDL, la dépense étant accrue par la distance entre les trois pôles d'implantation : Tours, Paris et Orléans.

Concernant les financements, les indications pertinentes figurent dans la partie 4.2 à laquelle nous renvoyons.

5.4 Valorisation

Les travaux du LLL sont orientés en linguistique fondamentale et didactique, avec un intérêt spécifique pour la description des langues, en sorte que les applications industrielles ne constituent pas un domaine d'application évident. Néanmoins, trois types d'études ont des implications qui pourraient appeler un certain développement au cours du prochain quadriennal, en lien avec l'informatique, la didactique et la psychologie :

Didacticiels

Par le lien entre linguistique, TAL et didactique, le LLL est le partenaire d'un Master Linguistique et didactique où a été mise en place une « niche » technologique concernant les didacticiels. Devenu un centre reconnu de formation dans ce domaine, comme le montrent les stages effectués et le placement professionnel des étudiants, en fonction des investissements matériels et humains de l'université, et en lien avec l'IUFM Centre-Val de Loire, de nouvelles applications sont envisageables qui constitueraient une équipe de recherche technologique à l'interface des TIC et de l'enseignement.

Ingénierie du document sonore

Dans la collaboration avec les équipes de la BnF et en profitant de l'expérience accumulée dans le domaine du traitement du document sonore, en particulier pour sa conservation, son catalogage/codage et sa disponibilité sur Internet, le LLL constitue un centre de ressources et d'expertise pour les questions des archives non graphiques, que cette compétence soit appelée pour du transfert de technologie (en particulier à destination de pays où la culture orale est une composante essentielle de la transmission patrimoniale) ou pour du conseil dans l'archivage des données.

Traitement Automatique des Langues

Dans le domaine des industries de la langue, un marché est en cours d'émergence dans lequel des chercheurs du LLL ont déjà pu signaler leur compétence, que ce soit en analyse du discours et détection de plagiat (N. Garric), dans l'implication des études cognitives dans le génie logiciel, dans les études sur l'automatisation des transcriptions (L. Hriba) ou dans l'accueil par des entreprises d'anciens étudiants de Master pour les études sur corpus oraux (Vecsys), l'extraction d'information (Sinequa, Lingway, Temis, LexisNexis) ou la traduction automatique (Lingua-et-Machina). D'autres applications sont envisageables à la condition de trouver un partenariat d'entreprise fiable qui fait encore défaut en région.

Sur ce point, une volonté politique, régionale ou nationale, qui contribuerait à la structuration d'un domaine dispersé entre deux instituts CNRS, les entreprises et les universités, serait un atout. La collaboration avec le L.I. à Tours (comme avec le LIMSI de Paris XI) ouvre des perspectives que la restructuration en Institut pourrait faciliter à Orléans dans un rapprochement en cours avec le LIFO, en apportant une nouvelle dynamique à des collaborations transversales.

5.5 Diffusion

La diffusion a toujours été au centre des préoccupations du LLL, ne serait-ce que par l'investissement de tous ses chercheurs dans l'enseignement, et plus encore à la BnF dont c'est la vocation première. S'il ne saurait être question de reprendre ici l'ensemble des missions de la BnF, on reviendra sur quelques initiatives menées dans le LLL, ou dans les laboratoires qui ont précédé sa constitution par fusion, et qui démontrent son implication particulière dans ce domaine.

Diffusion de revues et d'ouvrages

Dans le cadre du CERLICO, des collègues se sont investis depuis des années pour animer la revue en ligne CORELA (cf. présentation et annexes) qui associe numéros thématiques, articles hors thème et suppléments. Cette revue, bien référencée, est un relais important pour le CERLICO dont les actes de colloque sont publiés aux PUR dans l'année suivant leur tenue.

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

La *Revue de Sémantique et Pragmatique*, après des difficultés liées au service de publication de l'Université, a pu prendre un nouveau départ et assure aujourd'hui une représentation en langue française de ce domaine. Les sommaires des numéros sont joints en annexe, démontrant la diversité des écoles et la productivité du champ. Cette revue prolonge son activité par des « Rencontres de Sémantique et Pragmatique » dont la dernière édition a eu lieu à Gabès (Tunisie) en 2009.

Dans le domaine de l'hispanisme, la revue PROHEMIO a marqué sa singularité en investissant le domaine du lien entre culture écrite et culture orale, une approche aujourd'hui poursuivie, en lien avec les activités du laboratoire, sur le témoignage oral.

Enfin, il faut mentionner l'édition de la collection de l'« Encyclopédie des Langues et Cultures de Guyane » en co-édition CTHS-PUO dont trois volumes sont déjà parus (voir annexe). L'ensemble de ces supports ont été intégrés dans le programme du LLL pour le prochain quadriennal.

Diffusion électronique

Anticipant le rapprochement avec la BnF, une attention particulière a été consacrée à la diffusion des résultats de la recherche par voie électronique. Si, par manque d'IE, cette part du travail n'a pu aboutir à tous les résultats escomptés, notamment en matière de site du laboratoire, grâce au contrat passé avec la société ARES le LLL dispose à présent du plus grand corpus de français disponible en ligne, qui corresponde à la fois aux normes actuelles de ces produits et qui innove en particulier par son ergonomie et l'alignement des transcriptions sur le signal ou l'accès aux métadonnées. Cette expérimentation, conduite par O. Baude et C. Dugua a permis de tester les difficultés que soulève ce qui doit se poursuivre par une application aux corpus sonores de la BnF.

Comme un premier exemple de ces ressources, des documents extraits d'ESLO figurent déjà sur le portail de la DGLFLF (divers extraits d'entretiens illustrant le français central) et sur le site du journal *Libération* (extraits consacrés à Mai 68 vu d'Orléans).

Une autre création d'ampleur a été, à l'initiative d'E. Schang, la création du site de l'Association des Créoles à Base Lexicale Portugaise et Espagnole qui réunit des chercheurs dispersés sur quatre continents et facilite la disponibilité des données dans les pays du Sud.

Vulgarisation

La BnF, qui a toujours eu le souci d'assurer, par des expositions et des conférences, par la bibliothèque numérique Gallica, l'accès du plus grand nombre à ses ressources et à ses trésors, mène en ce domaine une activité particulièrement dynamique. Les chercheurs du LLL participent de leur côté à différentes actions de vulgarisation et assurent conférences, interviews et articles. On mentionnera le rôle de rédacteur d'O. Baude dans la revue du Ministère de la Culture et de la Communication ou l'article dans *Pour la science* de Mélanie Petit qui met à disposition du lectorat les résultats de sa thèse soutenue en 2009.

Ce volet d'activité revêt une importance particulière dès lors que la didactique, en particulier sous format numérique, est une activité importante dans les missions du laboratoire, en lien avec l'enseignement.

BIBLIOGRAPHIE CUMULEE

La décision de structurer autour d'un pivot l'ensemble des recherches a conduit à privilégier une présentation cumulée des productions. A l'exception d'ESLO, les axes et les équipes réunissent un petit nombre de chercheurs dont il est aisé de retrouver les productions dans les fiches individuelles récapitulatives. On a porté en gras les chercheurs qui étaient dans le laboratoire au moment de la publication (ou de la présentation en colloque).

Suivant les recommandations des associations professionnelles, nous n'avons pas fait de distinction entre les revues selon qu'elles sont ou non répertoriées dans des bases de données internationales. L'examen des dossiers étant réalisé par des collègues, nous considérons qu'ils savent faire la part des choses.

L'inventaire a respecté en revanche les classifications préconisées par l'AERES en pondérant, pour les publications en ligne, entre les différentes catégories, en fonction d'un critère de pertinence fondé sur la similitude des objets et du processus de publication. A été ajoutée une catégorie pour les interventions lors de journées d'études.

ACL : Articles dans des revues internationales ou nationales avec comité de lecture répertoriées dans les bases de données internationales

ACLN : Articles dans des revues avec comité de lecture non répertoriées dans des bases de données internationales

ASCL : Articles dans des revues sans comité de lecture

INV : Conférences données l'invitation du Comité d'organisation dans un congrès national ou international

ACT : Communications avec actes dans un congrès international ou national

COM : Communications orales sans actes dans un congrès international ou national

AFF : Communications par affiche dans un congrès international ou national

OS : Ouvrages scientifiques (ou chapitres de ces ouvrages)

OV : Ouvrages de vulgarisation (ou chapitres de ces ouvrages)

DO : Directions d'ouvrages

AP : Autres publications

TH : Thèses soutenues

Valorisation et brevets

Autres formes de production (brevets, ouvrages, rapports d'expertise)

ACL + ACLN

Articles dans des revues internationales ou nationales avec comité de lecture répertoriées dans les bases de données internationales &

Articles dans des revues avec comité de lecture non répertoriées dans des bases de données internationales

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>

Abouda, L. & Baude, O. (2009). « Du français fondamental aux ESLO », *Les Cahiers de Linguistique de Louvain*, 33, 2 : 131-146.

Ballier, N, & **Fournier, J.-M.** (2007), « La compétence phonologique en anglais au-delà des questions à l'agrégation », *Les Langues Modernes* 3/2007, Paris : APLV : XX-XX.

Bergounioux, G. (2006) « Indo-européen : de l'adjectif au nom », *Critique* 715 : 1004-1014.

Bergounioux, G. (2007) « La stylistique : Bally auteur du *Cours de Linguistique Générale* », *Pratiques*, 15-136 : 63-73.

Bergounioux, G. (à par.) « L'enjeu de l'apophonie dans les études indo-européennes », *Cahiers Ferdinand de Saussure* 62 (daté 2009).

Bergounioux, G. (à par.) « Saussure et l'histoire de la linguistique : l'apport des sources manuscrites », in « L'apport des manuscrits de Saussure », L. Depecker éd., *Langage* 178.

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.

Cadiot, P. (2006), « Métaphore prédicative nominale : genèse de la constitution de l'objet dans le champ subjectif », in *Aux carrefours du sens*, hommages offerts à Georges Kleiber pour son 60^e anniversaire, Leuven-Paris, Peeters : 641-657.

Cadiot, P. (2006) « Polysémie, discours et motifs lexicaux », in *RSP* 20 : 51-74

Cadiot, P. & Sanchez, J. (2007) « le cas des substantifs de couleur construits $X_A / \text{suf. eur} / X_A$ conv. X_n : un point de vue sur la synonymie », *Le Français Moderne*, 1 : 114-140

Cadiot, P. (2007) « A propos de l'anaphore connectivo-argumentative *pour le coup* », *Modèles linguistiques*, XXVIII-1 : 55-77

Cadiot, P. (2007) « Le morphème *pour* : polycatégorialité et unification sémantique », *Cahiers de lexicologie*, 90-1 : 75-91

Cadiot, P. & Visetti, Y.-M. (2008) « Proverbes, sens commun et communauté de langage », *Langages* 170 : 79-91

Cadiot, P. & Talmenssour, A. (2008) « Proverbes berbères et motifs proverbiaux » *RSP* 23 : 55-74.

Cadiot, P. (2009), « Drôle de drupe », en hommage à Claude Vandeloise, *CORELA*, téléchargeable sur :
<http://edel.univ-poitiers.fr/corela/document.php?id=2456>

Cadiot, P. (2009b), « Synonymie et couleur des mots », *Pratiques* : 24-44

Cambourian, A. (2006) « La pièce fait vide. Das Zimmer wirkt leer. Plaidoyer pour un petit système clos de l'apparence en allemand », *CORELA*,
URL : <http://edel.univ-poitiers.fr/corela/document.php?id=807>

Col, G. (2007) « Prosodie et émergence du sens. Propositions pour une étude cognitive de l'intonation » *Canadian Journal of Linguistics / Revue Canadienne de Linguistique*, 52(3) : 255-277.

Col, G. (2006) « *Appear, seem et look* : “ perception ” et “ construction ” des apparences » *CORELA — Les Verbes d'apparence*.
URL : <http://edel.univ-poitiers.fr/corela/document.php?id=813>

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. (2008) « Relation ou intégration prédicative ? La prédication comme principe d'émergence du sens », *Faits de Langues*, 31/32 : 33-43.

Col, G., Aptekman, J., Girault, S. et Victorri, B. (à par.) « Compositionnalité gestaltiste et construction du sens par instructions dynamiques », *CogniTextes*.

Col, G. (2010) « Des relations fonctionnelles des unités linguistiques aux relations instructionnelles », in *Espace, préposition, cognition - Hommage à Claude Vandeloise*, G. Col et C. Collin (eds). Numéro spécial de la revue *CORELA*.
URL : <http://edel.univ-poitiers.fr/corela/>

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.

Cristinoi, A. (2009) « Les erreurs dans la traduction automatique du genre dans les couples français-anglais et anglais-français : typologie, causes linguistiques et solutions », *Revue Française de Linguistique Appliquée*, XIV-1 : 93-108.

Chevrot, J.-P., Chabanal, D. & Dugua, C. (2007) « Pour un modèle de l'acquisition des liaisons basé sur l'usage : trois études de cas », *Journal of French Language Studies* 17 : 103-128.

Chevrot, J.-P., Dugua, C. & Fayol, M. (2009) « Liaison, word segmentation and construction in French: a usage-based account », *Journal of Child Language* 36 (3) : 557-596.

Delplanque, A. (2006) « Dérivation et production des adjectifs en dagara », *Cahiers voltaïques* 7, éd. K. Winkelmann & Manfred von Rocardor, Bayreuth : XX-XX.

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.

Dugua, C. (2008) « “un nouns”/“des nouns” ou le rôle de la fréquence sur l’acquisition de la liaison en français ». In Loiseau *et al.* (Eds.). *Autour des langues et du langage : perspective pluridisciplinaire*, Papiers sélectionnés du Colloque international des étudiants chercheurs en didactique des langues et en linguistique, Grenoble : PUG : 201-208.

Eshkol, I., Le Pesant D., (2007) « Trois petites études sur les prédicats de communication verbaux et nominaux », *Langue Française* 153 : 20-32.

Fabre, G. (2009) « Sérialisation 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>.

Chaboche, J. & **Fournié-Chaboche, S.** (à par.) « Club sportif professionnel et globalisation. Les dimensions, culturelles, discursives et territoriales de production et de marchandisation d’un mythe moderne : le cas du FC Barcelone », *Journal des Anthropologues* 120-121 (article de 45 000 signes, accepté et sous presse).

Gallèpe, Th. (2006) Anthroponymes en textes de théâtre : drôles de noms propres. *Méta* 51, 4 « *La traduction des noms propres* » : 651 - 59.

Gallèpe, Th. (2006) « Didascalies internes et construction de la représentation : l'exemple de *Napoleon oder die hundert Tage* de Chr. D. Grabbe », *Nouveaux Cahiers d'Allemand* 1 : 21 - 35.

Gallèpe, Th. (2007) « Anthroponymes, toponymes et autres « magiconymes » : leur traduction dans les versions française et allemande de *Harry Potter and the Order of the Phoenix* », *Nouveaux Cahiers d'Allemand* 4 : 351 – 368.

Gallèpe, Th. (2007) « *Gorbach mit Alois aus der Hütte* ou Didascalies sans verbes : types et fonctions », *Nouveaux Cahiers d'Allemand* 3 : 239 - 261.

Gallèpe, Th. (2009) « La "traduction" en allemand des noms propres en fiction narrative : jeux sémiologiques, génériques et textuels », *Cahiers d'études germaniques* 56, « Traduire, adapter, transposer » : 77 - 89.

Gallot, S., Spinelli, E., Chevrot, J.-P. & **Dugua, C.** (2009) « Le rattachement lexical de la liaison : une expérience d'amorçage chez des enfants pré-lecteurs », *Psychologie française* 54 : 355-361.

Garric N., Léglise I. & Point S. (2007) « Le rapport RSE, outils de légitimation ? Le cas TOTAL à la lumière d’une analyse de discours », *Revue de l'Organisation Responsable* 1, 2 : 5-19.

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

Garric N. & Longhi J. (2009) « Présentation », *Les Cahiers du LRL* 3, Université de Clermont-Ferrand : 11-24.

Garric N. (2009) « la médiation, entre connaissance de la discursivité et pratique discursive », *Les Cahiers du LRL* 3, Université de Clermont-Ferrand : 93-114.

Garric N. & Capedeвиelle-Mougnibas V. (2009) « La variation comme principe d’exploration d’entretiens : intérêts et limites de l’analyse lexicométrique interdisciplinaire pour l’étude de discours », *Corpus* 8 : 105-128.

Garric N. & Calas F. (2009) « La définition comme procédé discursif satirique dans les *Lettres persanes* de Montesquieu (1721) »

URL : http://www.publifarum.farum.it/publifarum_issues.php.

Garric N. & Goldberg M., (à par.), « Mise en scène de la scientificité dans le débat citoyen », *MEI (Médiation & information)* 31.

Gatelais, S. (2010) « *With* est-elle une préposition spatiale ? », in C. Delmas (ed.), *Faits de langues* 34 « Espace-temps anglais », Paris : Ophrys : 109-121.

Guérin, E. (2007) « ‘Moi je comprends pas et ça ça m’intrigue’. Réflexions sur la double occurrence de *ça* dans les énoncés oraux », *Linx*, 57 : 27-36.

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.

Guérin, E. (2009) « L’enseignement du FLM et la prise en compte de la nécessaire variabilité de la langue », *le français aujourd’hui*, 165 : 111-120.

Guérin, E. (à par.) « L’‘outre-langue’ des enseignants ou le mythe d’une langue monovariétale », *Pratiques*.

Hamma, B. (2006) « La préposition *par*, génératrice de polylexicalités », *LINX* 54, *Le semi-figement*, Vaguer, C. & Balibar-Mrabti, A. eds : 87-103.

Hamma, B. (2006) « Etat des lieux sur la sémantique de la préposition *par* », *Modèles linguistiques* 54, XXVII-2, *La préposition en français II* : 81-95.

Hamma, B. (2007) « La notion de ‘contraste’ dans les locutions de type *par N* », *Modèles Linguistiques* 55, XXVIII : 1, XX-XX

Hamma, B. (à par.) « De quelques ‘métaphores spatiales’ se construisant avec la préposition *par* », *RSP* 28.

Iglesias Ovejero, A.(2007) « Nombres del dominio francófono en el Madrid multicultural (1996-2006) », in Consuelo García Gallarín, coord., *Los nombres del Madrid multicultural*, Madrid, Ed. Partenón : 235-253.

Kihm, A. & **Rougé J.-L.** (à par.) « *Língua de Preto*, the Basic Variety at the root of West African Portuguese Creoles : A contribution to the theory of pidgin/creole formation as second-language acquisition » in *Journal of Pidgin and Creole Language*, Amsterdam : John Benjamins ed.

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

Lowrey, B. & **Toupin, F.** (2009) « La nature dynamique des langues, le changement et la notion d’invariant » *Bulletin des Anglicistes Médiévistes* 76 : 69-94.

Manuélian, H. & **Schang, E.** (2007) « XML, DTD et TEI pour un dictionnaire étymologique des créoles », *Dicorevue* oct. 2007, J. Pruvost Editeur : XX-XX.

Maurel D., Friburger N. & **Eshkol I.** (2009), « Who are you, you who speak? Transducer cascades for information retrieval », *4th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*, Poznań, 6-8 novembre, 220-223.

Nemo, F. (2006) « Contre la modularité ». *RSP* 19-20 : 27-50.

Nemo, F. (à par.). « Profilage temporel dans l'interprétation des morphèmes : de *toujours* à *tout* » *RSP* 25-26.

Nemo, F. (à par.) « Routines interprétatives, constructions grammaticales et constructions discursives » *Revue Estudos Linguísticos / Linguistic Studies*. Université de Lisboa.

Osu, S.N. & Clements, N. (2009) « Les nasales en ikwere », *CORELA* 7, 1.
URL : <http://edel.univ-poitiers.fr/corela/>

Osu, S.N. (2008) « La nature de la relation entre propositions dans une construction de type parataxe en ikwere », in Caron, B. (éd.) *Dépendance et parataxe dans les langues africaines*. Paris, Louvain, Peeters : 183-200.

Perrot, M.-E. (2006) « Statut et fonction du vernaculaire chiac : analyse de discours épilinguistiques », « Cultures minoritaires et urbanité : exploration, théories, méthodes », *Francophonies d'Amérique*, 22 : XX-XX.

Petit, M. (à par.) « Discrimination prosodique et représentation du lexique : les connecteurs discursifs », *Etudes de linguistique appliquée*.

Petit, M. (à par.), « Le marquage prosodique du commentaire énonciatif dans la discrimination du sens des mots de discours : l'exemple de *enfin* », *Travaux Neuchâtelois de Linguistique*.

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.

Rougé, J.-L. (2008) « Un dictionnaire des créoles portugais d'Afrique », *Dicorevue*, J. Pruvost éditeur.
URL : http://www.dicorevue.fr/bilingues/10-07_deca_4.html

Rougé J.-L. (à par.) : « Parler créole à Ziguinchor au 21^e siècle » *Sciences et techniques du langage*, Dakar, CLAD.

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

Toupin, F. (2010) « Exploring continuities and discontinuities between Ælfric's *Grammar* and its antique sources », *Neophilologus* volume 94, numéro 2 : 333-XX.
URL : <http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s11061-009-9185-5>

Toupin, F. (2008) « L'emploi hypothétique de *and* dans *Everyman* et en moyen-anglais », *Anglophonia* 24 : 23-44.

Toupin, F. (2006) « Les sens procéduraux intersubjectifs de la particule UP », *CORELA*, IV, 2.
URL : <http://edel.univ-poitiers.fr/corela/document.php?id=1355>

Vlad, D. (2007), « Du discours polémique à la polémique dans le discours », *Revue de Sémantique et Pragmatique* 21/22 : 82-94.

Vlad, D. (2010), « Degrés de plurivocité et formes de polyphonie », B. Lépinette, B. Gómez Ángel (éds.), *Etudes de linguistique française, Quaderns de filologia de la Universitat de València*, 71 : 231-239.

Vlad, D. (2010), « La polémique – une forme particulière de communication conflictuelle », *Studia Universitatis Babeş-Bolyai, Philologia* 1 : XX-XX.

Wolf-Mandroux, Fr. (2008) « Teaching and discovering a culture and a language by the creation and the use of games: an experience of French primary trainee teachers », Colloque international, Shaker Verlag, Aix-la-Chapelle, « Potsdamer Beiträge zu Sprach- und Kulturmittlung » 1 : XX-XX.

ASCL

Articles dans des revues sans comité de lecture

Bergounioux, G. (2007) « Epistémologie de la linguistique : Analyse, comparaison, structure et typologie »

URL : <http://www.accedit.com/auteur.php?id=bergounioux>

Bergounioux, G. (2007) « Un art poétique aujourd'hui ? A partir de poèmes de L. Boltanski, de J.-P. Michel et J. Sacré »

URL : <http://www.accedit.com/auteur.php?id=bergounioux>

Bergounioux, G. (2008) « Le langage des enfants dans la pensée française du XIX^e siècle »

URL : <http://www.accedit.com/auteur.php?id=bergounioux>

Bergounioux, G. (2008) « L'histoire sociale de la linguistique et la sémantique »

URL : <http://www.accedit.com/auteur.php?id=bergounioux>

Bourdier, Ph. (2006) « La science des bulles », *Au clair de la plume : le documentaire scientifique*, 80, Chartres, ACP : 18-24.

Bourdier, Ph. (2008) « Promenade rétivienne dans Paris », *Etudes Rétiviennes*, 40 : 174-175.

Bourdier, Ph. (2010) « L'enseignement du cinéma dans les années 50 en France : l'exemple des actions de *Film et jeunesse* », *Images à l'école, image de l'école, 1880-1960*, éd. Musée de l'école : 35-45.

Cadiot, P. (2006) « Le signifié : un objet non identifié », *Cahiers de linguistique analogique* 2 : 27-44.

Dugua, C. & Spinelli, E. (2009) « La liaison : Effets de la fréquence et du rapport à l'écrit sur son acquisition et son usage », *Les cahiers de l'infantile* 7 : 17-41.

Dugua, C., Chevrot, J.-P. & Fayol, M. (2006) « Liaison, segmentation des mots et schémas syntaxiques entre 2 et 6 ans : un scénario développemental », *Entretiens de Bichat : Orthophonie*, Paris : Expansion, formation et éditions : 230-244.

Fournié-Chaboche, S. (2006) « Las palabras españolas de la lengua francesa: aspectos semánticos de la transferencia cultural », *Cahiers du P.R.O.H.E.M.I.O.* 7 (*Interferencias lingüísticas y culturales en El Rebollar y otras partes*) : 261-279.

Hamma, B. (2010) « La non-compositionnalité est-elle un indice du caractère locutionnel? Etude de quelques locutions en *sous* », in *Les Cahiers de l'ED 139, Syntaxe et Sémantique grammaticale Descriptives*, Publication de Paris Ouest Nanterre-la Défense : 55-86.

Hamma, B. (à par.) « Vers une redéfinition des définitions lexicographiques. Remarques sur les signifiés de *beau* et *dent* », *Revue de Lexicologie*, Tunis-la Manouba.

Iglesias Ovejero, A. (2006) « Avatares lingüísticos del Águeda, un río con nombre de mujer : Apuntes para un esbozo de hidronimia serragatina », *Ciudad Rodrigo, Carnaval*, 2006 : 425-438.

Iglesias Ovejero, A. (2008) « La feria de *Guinaldu* desde la nostalgia infantil », *Fuenteguinaldo, Fiestas 2008* : 157-159.

Iglesias Ovejero, A. (2008) « Santos y diablos en los pagos del Carnaval », *Ciudad Rodrigo, Carnaval*, 2008 : 317-324.

Iglesias Ovejero, A. (2009) « ¿Cualquiera tiempo pasado fue mejor?: Fuenteguinaldo y sus aledaños del siglo XVIII al XX a ojo de buen cubero », *Fuenteguinaldo, Fiestas 2009* : 170-174.

Iglesias Ovejero, A. (2009) « Los anecdotarios de bravuconadas y la gran mojaína de 1807 », *Ciudad Rodrigo, Carnaval*, 2009 : 77-83.

Motei, A. (2009) «Traces de la liberté surveillée dans les reportages», *Interstudia. Langue et Société, Discours et Pouvoir*, Bacau, 4 : 82-88.

Motei, A. (2008) «Manipularea în presa scrisă – între inevitabil și intenționat » [La manipulation dans la presse écrite – entre inévitable et intentionné], *Studia Universitatis « Babeș-Bolyai », Ephemerides*, Cluj-Napoca, 2 : 41-52.

Motei, A. (2007) «Sélection des événements et couvertures thématiques dans trois quotidiens roumains d'information générale», *Studia Universitatis « Babeș-Bolyai », Ephemerides*, Cluj-Napoca, 2 : 65-77.

Nemo, F. (2006) « Eléments pour une typologie linguistique des rapports forme/sens », *Cahier de linguistique analogique* : 205-226.

Osu, S.N. (2008) « Verb reduplication in Ikwere: beyond form and meaning », *Revue Gabonaise des Sciences du Langage* 3-4 : 5-36.

Wolf-Mandroux, Fr. (2006), « The DCL in France: a new diploma based on the Common European Framework of Reference for Languages », in *Neos*, 1, « Journal of New Trends in Education », *Revue Universitaire de Cakovec, Croatie* : XX-XX

INV

Conférences données à l'invitation du Comité d'organisation dans un congrès national ou international

Baude O. (2006) « Pierre Encrevé et la réforme de l'orthographe : «le champ du linguiste », colloque international *Faire signe*, 16-18 octobre, Paris.

Baude, O. (2007) « Contributions des corpus oraux à la linguistique de corpus : *une démarche réflexive intégrée* » 5^e Journées de la Linguistique de Corpus, Lorient, 13 – 15 septembre.

Bergounioux, G. (2009) « La fonction critique de l'histoire de la linguistique » « Discours sur les langues et rêves identitaires », Velmezova & Sériot ed. Lausanne, *Cahiers de l'ILSL* 26 : 5-19.

Bergounioux, G. (à par.) « Du corpus en linguistique : les deux paradigmes du structuralisme », actes du colloque « Pour une épistémologie de la sociolinguistique », Montpellier, 10-12 décembre 2009, H. Boyer éd.

Cristinoi, A. (2009) « Traduction vs traduction automatique – un conflit insoluble ? », Table ronde du Colloque Sciences du langage et Nouvelles technologies, Paris, 5 décembre 2009.

Dugua, C., Chevrot, J.-P. & Fayol, M. (2006) « Liaison, segmentation des mots et schémas syntaxiques entre 2 et 6 ans : un scénario développemental » Conférence invitée aux entretiens de Bichat : Orthophonie, Paris, 29-30 septembre 2006.

Gallèpe, Th. (2008) Conférence du Consulat Général de France à Hambourg et de l'ICN pour la présentation du livre « Théorie du récit – L'apport de la narratologie allemande » édité par J. Pier (Presses Universitaires du Septentrion) intitulée : « Narratologie übersetzen ».

URL : http://www.icn.uni-hamburg.de/webfm_send/59

Gallèpe, Th. (2008) Conférence au « Workshop » organisé par l'ICN (Interdisziplinäres Centrum für Narratologie der Universität Hamburg) intitulée : « Stanzel au pays de Genette »

URL : http://www.icn.uni-hamburg.de/webfm_send/58

Garric N. (2010) « Maîtriser l'hétérogénéité des corpus par la variation des corpus, des outils et des méthodes », conférence à la Journée d'étude « L'analyse de corpus face à l'hétérogénéité des données », 6 mai 2010, Paris-INA.

Grenand, F. (2006) « Très « corações » sozinhos... ou como (des-)construir um parque nacional na Guiana francesa », conférence au colloque Guiana Amerindia : etnologia e história, Museu Paraense Emílio Goeldi , Belém do Para, Brésil, 31 octobre-2 novembre 2006.

Grenand, F. (2006) « Em busca da língua omawense : uma pesquisa quase arqueológica », conférence au colloque Novidades sobre o Homem Americano, Museu do Homem Americano, São Raimundo Nonato, Piauí, Brésil, 14-16 décembre 2006.

Grenand, F. (2008) « Observatório sociedades / meio ambiente: O Oiapoque, um rio compartilhado », conférence au Secundo Simpósio Regional de Farmacognosia, Pirenópolis, Goiás, Brésil, 16-19 octobre 2008.

Rougé, J.-L. (2010) « A natureza dos crioulos » Conférence aux 6^e rencontres de l'ABECS (Associação Brasileira de Estudos dos Crioulos e Similares) Universidade Federal da Bahia (Salvador da Bahia – Brésil), Juin.

ACT

Communications avec actes dans un congrès international ou national

Abasq, V. (à par.) « Préfixe ex- et accentuation », in *Actes du XIIIe Colloque d'Anglais Oral de l'ALOES*, Paris XIII- Villetaneuse 1^{er} avril 2008.

Abasq, V. (à par.) « Comportement accentuel des substantifs en out-, over- et under- », in *Actes du XIVe Colloque d'Anglais Oral de l'ALOES*, Paris XIII- Villetaneuse 4 avril 2008.

Abouda, L. (2010) « De la visibilité syntaxique des modes, de l'invisibilité syntaxique des temps », in Álvarez Castro, C., Bango de la Campa, F. & Donaire, M. L. (eds), *Liens linguistiques. Etudes sur la combinatoire et la hiérarchie des composants*, Collection « Sciences pour la Communication », Berne, Peter Lang, Vol. 90 : 319-334.

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^e Colloque d'Albi, Toulouse, Presses Universitaires du Mirail : 161-168.

Baude, O. (2007) « Corpus oraux les bonnes pratiques d'une communauté scientifique », *Corpus en Lettres et Sciences sociales. Des documents numériques à l'interprétation*, Rastier, F. & Ballabriga M. dir, Actes du XXVII^e Colloque d'Albi, Toulouse, Presses Universitaires du Mirail : 61-66.

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.

Bellonie, J.-D. & **Guérin, E.** (à par.) « Lorsque la réflexion sociolinguistique éclaire la problématique de l'enseignement du FLM », Actes du colloque international *Pour une épistémologie de la sociolinguistique*, Montpellier, 10-12 décembre 2009.

Bergounioux, G. (2006) « L'endophasie dans la théorie des opérations énonciatives », actes du colloque de Cerisy-la-Salle, in *Antoine Culioli, un homme dans le langage* (D. Ducard et Cl. Normand eds) : 101-116.

Bergounioux, G. (2006) « Vers le *Mémoire* ou comment le structuralisme vint à Saussure », actes du colloque de la Société d'Histoire et d'Épistémologie des Sciences du Langage, « Aux origines du structuralisme en linguistique » (ENS Lyon, janvier 2006), C. Puech, éd.

URL : <http://htl.linguist.jussieu.fr/num3.htm/bergou.pdf>

Bergounioux, G., Bergounioux M., Nguyen N. & Wauquier S. eds (2007) « Quel programme de recherche en mathématiques et phonologie » in G. Bergounioux *et al.* (2007) : 9-26.

Bergounioux, G. (2007) « From a reference corpus to a prototypical corpus : ESLO1 & ESLO2 » *Corpus Linguistics*, Birmingham, Actes du colloque *Corpus Linguistics*, Birmingham. Paper 64 : 3-9

URL : http://www.corpus.bham.ac.uk/corplingproceedings07/paper/64_Paper.pdf

Bergounioux, G. (2008) « Des œufs mués : entendre omelette » Actes du 1^{er} *Congrès Mondial de Linguistique Française* organisé par PILF-CNRS, Paris, 9-12 juillet 2008), en CD-ROM et

URL : <http://www.linguistiquefrancaise.org/>

Bergounioux, G. (2010) « Mai 68 vu d'Orléans », Actes du 2^e 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/>

Bergounioux, G. (2010) « La phonologie comme morphologie » *Le Projet de Ferdinand de Saussure*, Bronckart, J.-P., Bulea, E. & Bota, C. (eds), Genève-Paris, Droz : 105-124.

Bergounioux, G. (à par.) « Entre folie et raison : la genèse du délire d'interprétation », actes du colloque « Langues imaginaires, imaginaire du langage » (Genève, 5-6 décembre 2008), Pot, O. éd., *Le Genre Humain*, Paris, Seuil.

Bourdier, Ph. (2007) *La notion de genre dans les manuels de français actuels : le stéréotypage du genre cinématographique*, Stéréotypes et stéréotypage, actes du colloque international de Montpellier, juin 2006, vol. 3, P.U.de Montpellier-L'Harmattan : 128-133.

Bourdier, Ph. (2008) « Prendre en considération les représentations et les pratiques des enseignants des professeurs des écoles à l'égard du cinéma dans les parcours de formation initiale », colloque *Eduquer aux médias, ça s'apprend*, Clermont-Ferrand, 6-7 décembre 2007, organisé par le Ministère de l'Education Nationale, l'IUFM d'Auvergne, le CLEMI, atelier 17

Chaboche, J. & **Fournié-Chaboche S.** (2009) « Entre *naming* et *naming rights*: les enjeux identitaires des stratégies de dénomination des stades en Europe », *Colloque international Sport et communication*, IDETCOM, In *PUSS : Annales de l'Université Toulouse 1-Capitole, Tome LI, XX-XX*.

Cavalheiro, M. (à par.) « Plurilinguisme et insécurité linguistique en milieu scolaire : le cas d'écoliers ouagalais, au Burkina Faso », Colloque International « La construction identitaire à l'école. Approches pluridisciplinaires ». IUFM de Montpellier (UM2), du 6 au 8 juillet 2009.

Cavalheiro, M. (à par.) « Représentations et constructions identitaires en milieu scolaire ouagalais (Burkina Faso) », Colloque International des Jeunes Chercheurs du CERCI (Centre de Recherche sur les Conflits d'Interprétation) « Sens et représentations en conflit ». Université de Nantes, du 3 au 5 décembre 2009.

Cavalheiro, M. (à par.) « Le français du Burkina Faso : norme scolaire et variations. L'exemple du discours rapporté dans des écrits d'écoliers ouagalais », *Actes du XXVI^e Congrès International de Linguistique et Philologie Romanes*, Université de Valence (Espagne).

Cavalheiro, M. (à par.) « Les usages déclaratifs d'écoliers ouagalais (Burkina Faso) : qu'est-ce qu'être "un bon francophone" ? », *Actes du XXVI^e Congrès International de Linguistique et Philologie Romanes*, Université de Valence (Espagne).

Chaboche, J. & **Fournié-Chaboche S.** (à par.) « *FC Barcelona : més que un club* : facteurs culturels et discursifs de la production et de la marchandisation d'un mythe sportif », in *Actes du Colloque international Mythologies et mondialisation*, Université de Nantes, 10 et 11 juin 2008.

Clavier, V. & **Lafont-Terranova, J.** (2006) « Le français dans une filière technologique : une approche transversale pour l'apprentissage des discours de spécialité », in *Actes du 9^e colloque international de l'Association internationale pour la recherche en didactique du Français (AIRDF), Le français : discipline singulière, plurielle ou transversale ?* Québec, 26-28 août 2004, CD-ROM ou URL : <http://www.colloqueairdf.fse.ulaval.ca/fichier/Communications/clavier-lafont-terranova.pdf>

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. (2009) « Activation et saillance d'unités grammaticales. Quelle place pour l'intonation dans la construction dynamique du sens ? » *Travaux linguistiques du Cerlico*, 22, PUR : 59-79.

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.

Col, G. (à par.) « Propositions pour une étude de l'intonation et de son rôle dans l'émergence du sens », *Actes du XI^e Colloque d'Anglais Oral de l'ALOES*, Paris XIII-Villetaneuse.

Col, G. (à par.) « Modèle instructionnel du rôle des unités linguistiques dans la construction dynamique du sens ». *Actes du Congrès International « L'Homme au cœur des Dynamiques Sociales, Territoriales et Culturelles »* (MSHS de Poitiers). Rennes : Presses Universitaires.

Col, G. (à par.) « Remarques sur le rôle de l'intonation dans le traitement de l'information. », *Actes de la Journée Scientifique « Parole »* (Poitiers), numéro spécial de la revue *CORELA*.

Col, G. (à par.) « Correspondance et mixage d'espaces mentaux dans la construction dynamique du sens. », *Mémoires de la Société Linguistique de Paris* 18. Louvain, Peeters.

Col, G. & Aptekman, J. (à par.) « Construction du sens par instructions dynamiques », *Rencontres Interdisciplinaires sur les Systèmes Complexes Naturels et Artificiels*, Rochebrune, Janvier 2010. (à par.)

Dugua, C. & Chabanal, D. (2006) « Acquisition de la liaison chez l'enfant francophone : formes lexicales de Mots2 », in *Actes des XXV^es Journées d'Etudes sur la Parole*, Dinard 12-16 juin 2006 : XX-XX.

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/>

Eshkol I., Tellier I., Taalab S. & Billot S. (à par.) « Étiqueter un corpus oral par apprentissage automatique à l'aide de connaissances linguistiques », *10th International Conference on statistical analysis of textual data (IADT 2010)*, Rome, 9-11 juin. *Actes en ligne*.

Eshkol I., (à par.) « Interpréter le contexte dans un corpus oral : fonctions et limites du traitement automatique des données linguistiques », in *Actes des Rencontres Interdisciplinaires sur les Systèmes Complexes Naturels et Artificiels.*, Rochebrune, France 17-24 janvier 2010.

Eshkol I., (à par.) « A quelles conditions une taxinomie du lexique est-elle possible en TAL ? », *Actes du colloque international : L'illusion taxinomique*, Tunis, Tunisie 19-21 novembre 2009.

Eshkol I., (2010) « Entrer dans l'anonymat. Etude des "entités dénommantes" dans un corpus oral », *Eigennamen in der gesprochenen Sprache*, Narr Francke Attempto Verlag : 245-266.

Fournié-Chaboche, S. (2010) « Monstres aimés et amours monstrueuses : analyse de la relation littéraire entre monstre et amour dans *Amado monstruo* de Javier Tomeo », in *Actes du Colloque international ALMOREAL « Amours interdites »*, P.U.O. : 221-233.

Fournié-Chaboche, S. (à par.) « L'écriture en spirale de Javier Tomeo : analyse et modélisation », *Actes du Colloque international ALMOREAL « Chemins de traverses »*, Université d'Angers, mars 2010

Fournier J.-M. (à par.), « Accent lexical et structure syllabique en anglais : l'accent préfère-t-il les agrégats consonantiques ? », 15^e colloque d'avril sur l'anglais oral, Université de Paris-Nord.

Gallèpe, Th. (2008) « De *Hinze und Kunze* au *Hinze-Kunze-Roman* : du vrai théâtre au faux roman ; métalepse et ironie In *Autour du Hinze-Kunze-Roman de Volker Braun – Parcours interprétatifs* ». Dijon: Individu et Nation (ISSN 1961-9731) du Centre Interlangues Textes Images Langages (EA 4182) de l'Université de Bourgogne :

Garric N., Capedevielle-Mougnibas I. & Besse M-O. (2006) « Intérêts et limites de l'analyse du discours pour la recherche interdisciplinaire et la coproduction de connaissances scientifiques. Le cas d'une analyse lexicométrique d'entretiens avec Lexico3 », *Actes des JADT 2006*, Besançon : 439-450.

Garric N. & Gazeau M-A. (2007) « *L'Hebdo du médiateur* : un espace de démocratie illusoire », in *Actes du Colloque international Les mises en scène du discours médiatique*, 21 au 23 juin 2007, Université de Laval, Québec :

URL : <http://www.com.ulaval.ca/lab-o/actes.php>.

Garric N. & Légèze 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.

Garric N., (à par.) « *L'Hebdo du médiateur* : construction et déconstruction d'une demande sociale », in *Actes du Colloque international Analyse du Discours de Demande Sociale : enjeux théoriques et méthodologiques* » (27-29 novembre 2008, Paris).

Grenand, F. (2008) « Des encyclopédies amérindiennes pour valoriser des patrimoines immatériels », in *Actes du colloque de l'Académie des Technologies : Développer et préserver : technologies, cultures, environnement et développement durable* (Cayenne 29 octobre-4 novembre 2007) : 33-34.

Guérin, E. (2008). « Le français standard : une variété située », in *Actes du I^{er} Congrès Mondial de Linguistique Française 08*, Paris, 9-12 juillet 2008.

<http://www.linguistiquefrancaise.org/>

Guérin, E. (2009) « Authenticité des données et description grammaticale scolaire », *Actes du colloque international Des documents authentiques oraux aux corpus : questions d'apprentissage en didactique des langues* (Nancy, 14-15 décembre 2007), *Mélanges-CRAPEL* 31 : 147-157.

Guérin, E. (à par.) « Le mythe de la dichotomie oral/écrit entretenu dans les manuels de grammaire à travers la présentation des temps du passé », in *Actes du colloque international « Grammaire en francophonie » - Curricula, manuels, pratiques d'enseignement observées, formation des enseignants*, Montpellier, 16-18 février 2010.

Guérin, E. (à par.) « « *C'est à l'école qu'on apprend à parler poli* ». Construire son identité linguistique étant donné l'idéologie du « bon usage » », in *Actes du colloque international La construction identitaire à l'école Approches pluridisciplinaires*, Montpellier, 6- 8 juillet 2009.

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.

Hamma, B. (2009) « La place de la pensée dans l'étude de la langue. Confrontations des thèses cognitivistes et de celles de Ferdinand de Saussure », (actes du colloque « Révolutions saussuriennes ») in *Du côté de chez Saussure*, M. Arrivé (éd.), Lambert-Lucas, Limoges : 133-150.

Hamma, B. (à par.) « *Esprit, tête, main* et autres 'lieux de passage métaphoriques' se construisant avec *par* », *V^e Rencontres de Sémantique et Pragmatique*, RSP.

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.

Iglesias Ovejero, A. & Giraud, F. (2006) « Cuán necios y cornudos son *los otros* en El Rebollar y otras partes: *Portuguesis, Serranus* y *Farinatus* según el sociocentrismo del refranero castellano, extremeño y rebollano » en collaboration avec Françoise Giraud, *Cahiers du PROHEMIO 8, Interferencias lingüísticas y culturales en El Rebollar y otras partes*. Actas de las II Jornadas Internacionales de Estudio de El Rebollar (VI Colloque PROHEMIO, 22-25 juillet 2005, P.U.O. : 363-400.

Iglesias Ovejero, A. & Giraud, F. (2006) « Tres relatos orales rebollanos: *El lobu y los pastoris, Los polvos de las cocas* y *La Muerti* », *Cahiers du PROHEMIO 8, Interferencias lingüísticas y culturales en El Rebollar y otras partes*. Actas de las II Jornadas Internacionales de Estudio de El Rebollar (VI Colloque PROHEMIO, 22-25 juillet 2005, P.U.O. : 757-766.

Iglesias Ovejero, A. (2006) « El juego de la traducción y Cervantes : Algunas observaciones sobre interferencias lingüísticas en el *Quijote* y su última traducción al francés (2001) », *Le premier Quichotte de 1605 et ses avatars littéraires et iconographiques* (Actes de la journée de Recherche sur le *Quichotte*, 14 octobre 2005, Université Lumière Lyon 2- Institut Cervantès), *Crisoladas*, 1, Ed. Crisol 16/17 : 11-52.

Iglesias Ovejero, A. (2008), « Archivos vivientes: las víctimas del terror militar de 1936 a 1939 en El Rebollar y pueblos aledaños salmantinos », in *Cahiers du PROHEMIO 9, 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 PROHEMIO, 20-22 juillet 2007 : 101-201.

Iglesias Ovejero, A. (2008) « Memorias del terror: Transcripción literal de testimonios de Robleda (R 1973, R 1976) y El Payo (EP 1973) », in *Cahiers du PROHEMIO 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 PROHEMIO, 20-22 juillet 2007 : 473-549.

Iglesias Ovejero, A. (2008) « La estrategia de los vencedores de 1936-1939 y su *hybris* en El Rebollar salmantino : la memoria de las víctimas », in *Mémoire(s). Représentations et transmission dans le monde hispanique (XX^e-XXI^e siècles)*, actes du colloque *Hispanística XX*, Dijon, 23-24 novembre 2007, Catherine Orsini-Saillet, ed., *Hispanística XX*, 25, Editions Universitaires de Dijon : 97-114.

Iglesias Ovejero, A. (2010) « Ensayo de cronología del alzamiento militar, terror y represión de 1936 a 1946 en el Alto Águeda y otras localidades de la tierra de Ciudad Rodrigo », à paraître (2d trimestre 2010) in *Cahiers du PROHEMIO 11, DVD, Historia, habla y cultura tradicional*. Actas de las IV Jornadas Internacionales de Estudio de El Rebollar (VIII Coloquio PROHEMIO, 17-19 juillet 2009, P.U.O. : 177-320.

Iglesias Ovejero, A. (2010) « Apotegmática : hechos y dichos tradicionales de El Rebollar », *Cahiers du PROHEMIO* 11, DVD, *Historia, habla y cultura tradicional*. Actas de las IV Jornadas Internacionales de Estudio de El Rebollar (VIII Coloquio PROHEMIO, 17-19 juillet 2009, P.U.O. : 381-410.

Iglesias Ovejero, A. (2010) « Apotegmática : hechos y dichos tradicionales de El Rebollar », *Cahiers du PROHEMIO* 11, DVD, *Historia, habla y cultura tradicional*. Actas de las IV Jornadas Internacionales de Estudio de El Rebollar (VIII Coloquio PROHEMIO, 17-19 juillet 2009, P.U.O. : 381-410.

Isidore-Prigent, J., (2006) « Quelles compétences pour étayer les raisonnements orthographiques des élèves ? » in *Former des enseignants professionnels*, Actes du colloque international Recherche(s) et Formation, IUFM Pays de Loire, février 2006.

Isidore-Prigent J. (2007) « Dialogues didactiques pour une lecture littéraire au cycle 3 : étude comparée de la conduite discursive de deux enseignants et de ses effets sur les élèves », in *Les effets des pratiques enseignantes sur les apprentissages des élèves*, Actes du colloque international PNE, IUFM Besançon, mars 2007.

Lafont-Terranova, J. & Colin, D. (2006) « L'écriture en français et dans deux autres disciplines : représentations d'enseignants », in *Actes du 9^e colloque international de l'Association internationale pour la recherche en didactique du français (AIRDF), Le français : discipline singulière, plurielle ou transversale ?* Québec, 26-28 août 2004), CD-ROM ou
URL : <http://www.colloqueairdf.fse.ulaval.ca/>

Lafont-Terranova, J. (2007) « Un atelier d'écriture créative dans l'enseignement supérieur technologique : analyse d'une expérience », in *Actes 5 du 2^e colloque international de littérature de l'Association belge pour la lecture (ABLF), Apprendre ensemble à (mieux) lire et écrire*, Namur 3 et 4 novembre 2005, *Caractères* 25 : 4-9.

Lafont-Terranova, J. (2008) « Travailler conjointement l'investissement de l'écriture et les conceptions de l'écriture du sujet-écrivain », symposium « Rapport à l'écrit/à l'écriture », *Dixièmes rencontres du Réseau REF*, Sherbrooke (Québec) 9 et 10 octobre 2007, in S.-G. Chartrand & C. Blaser (dir.), *Le rapport à l'écrit : un outil pour enseigner de l'école à l'université*, CEDOCEF /Presses universitaires de Namur (Diptyque 12) : 26-42.

Lafont-Terranova, J., Colin, D. & Isidore-Prigent, J. (2010) « Etude du rapport à l'écriture d'enseignants du primaire (CM2) et du secondaire (6^e) dans une perspective curriculaire », *XI^e 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.

Lafont-Terranova, J. & Niwese, M. (à par.) « De l'écriture créative à l'écriture de recherche : analyse d'une expérience de formation à la didactique de l'écriture », *colloque international Littéracies universitaires : Savoirs, écrits, disciplines*, Université Lille 3, 2-4 septembre 2010. Actes sur CD-ROM

Motei, A. (2007) « Les témoignages, un type de discours de presse passé sous silence », in *Actes du Colloque international de langue et littérature française « Continuité et modernité du français »*, Brasov, 24 - 26 mai 2007 : 111-124.

Motei, A. (2007) «Construction et organisation de la *Une* dans trois quotidiens roumains», in *Proceedings of the International Conference «Individual and Specific Signs. Paradigms of Identity in Managing Social Representations»* Bacău, 25 - 28 octobre 2007 : 433-441.

Motei, A. (2006) «L'éditorial – type de discours de presse entre les contraintes contractuelles et les choix individualisants des acteurs du monde médiatique», *Actes des VIIIèmes Rencontres Jeunes Chercheurs de l'École Doctorale 268 Langues et Langues, Paris Sorbonne Nouvelle, 21 mai 2005*, publiés en mai 2006 : 102-106.

Motei, A. (à par.) «Construction du thème journalistique», in *Actes du Colloque International de Linguistique - Le texte : modèles, méthodes, perspectives*, Cluj-Napoca 25 - 27 septembre 2008.

Osu, S.N. (2006) « Comment j'étudie une langue africaine au 21^e siècle », in *Rencontre du Réseau thématique prioritaire Afrique*. Paris, CNRS. Atelier : la linguistique africaine : sa place, ses enjeux au XXI^e siècle. Publication en ligne. Version **texte**.

URL : <http://www.etudes-africaines.cnrs.fr/communications/osu.pdf>

Osu, S.N. (2010) « Entre reduplication et répétition : de l'identification à l'identité », in Osu, S. N., Col, G., Garric, N., Toupin, F. (éds), *Construction d'identité et processus d'identification*. Berne, Peter Lang : 545-568.

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 .

Perrot, M.-E. (à par.) « Acadieman et l'Académie chiac : le chiac de l'oral à l'écrit », *Actes du colloque de l'AFLS Les voix du français* (Oxford, septembre 2008), Peter Lang.

Petit, M. (à par.) « Le rôle de la prosodie dans le traitement automatique du sens : l'exemple de *enfin* dans un corpus de français parlé », 6^e Journées Internationales de Linguistique de Corpus, Université de Bretagne Sud, 10-12 septembre 2009.

Petit, M. (à par.) « Rôle de la prosodie dans la caractérisation sémantique des unités lexicales », 5^e Rencontres de Sémantique et Pragmatique, Université de Gabès (Tunisie), 22-24 avril 2009.

Petit, M. (à par.) « L'intégration de l'information prosodique en lexicographie : nouveaux supports, formats de présentation et techniques de discrimination », Colloque International « Lexicographie et informatique : bilan et perspectives », Université de Nancy 2, 23-25 janvier 2008.

Rougé, J.-L. & Schang, E. (2010). « Ce que nous enseigne la comparaison des créoles portugais d'Afrique ». XXVI^e Congrès Internacional de Lingüística i Filologia Romàniques. Valencia, 5-11 septembre 2010.

Serpollet, N., Bergounioux, G., Chesneau, A. & Walter, R. (2007) « A Large Reference Corpus for Spoken French : ESLO 1 and 2 and Its Variations », article 64, 41 pages.

URL : <http://ucrel.lancs.ac.uk/publications/CL2007/>

Serpollet, N. (2007) « 'Tell me how you cook and I will tell you who you are – How can a question such as “How do you make an omelette?” illustrate the morpho-syntactic and sociological variations found in the ESLO oral corpus? », in Davies, M., Rayson, P., Hunston, S.

& Danielsson, P. (eds), *Proceedings of the Corpus Linguistics Conference, CL2007*, University of Birmingham, 27-30 juillet 2007 : 23-40.

URL : http://ucrel.lancs.ac.uk/publications/CL2007/paper/64_Paper.pdf

Toupin, F. (2006) « L'évolution sémantique de *up* à la lumière de la pragmatique historique », *Travaux du CIEREC* 128, Publications de l'Université de Saint-Etienne : 33-48.

Toupin, F. (2006) « Le Préverbe *be-* du vieil-anglais », *Travaux de Diachronie* 3, Presses universitaires François Rabelais : 59-79.

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.

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^e-XV^e siècles)*. Presses universitaires Paris-Sorbonne : 179-201.

Toupin, F. (2009) « Les propositions hypothétiques dans la moralité *Everyman* », *Everyman : actes des colloques de janvier 2009*, Presses de l'Université de Nancy : 95-111.

Vlad, D. (2006), « Sur quelques marqueurs polyphoniques à valeur polémique », in F. Hrubaru, A. Velicu (éds.), *Enonciation et syntaxe, Actes du XII^e Séminaire de Didactique Universitaire, Constanta 2005*, Echinox, Cluj-Napoca : 205-222.

Vlad, D. (2006), « Quand interroger, c'est contester », in R. Superceanu, D. Dejica (eds.) *Comunicare profesională și traductologie, Lucrările conferinței internaționale 29-30 septembrie 2005, Timișoara*, Editura Politehnica, Timișoara, 2006: 81-88.

Vlad, D. (2008), « Degrés de plurivocité et formes de polyphonie », in B. Lépinette, B. Gómez Ángel (éds.), *Actas 7 Congreso Internacional de Lingüística francesa « Linguistique plurielle », Valencia, 25-27 octobre 2006*, Editorial de la Universidad Politécnica de Valencia, Valencia, 2008: 279-286 (CD-Rom).

Vlad, D. (à par.) « La polémique – phénomène discursif transgénérique », in *Actes du XXV^e Congrès International de Linguistique et Philologie Romanes, 2-8 septembre 2007, Innsbruck*.

Vlad, D. (à par.) « Le texte polémique : confrontation dialogique de discours dans le monologal », in *Actes du colloque International « Le texte : modèles, méthodes, perspectives », Universitatea « Babeș Bolyai » Cluj-Napoca, 24-27 septembre 2008*.

Vlad, D. (à par.) « Pour une typologie des formes de communication conflictuelle », in *Actes du colloque « Journées de la francophonie », XIV^e édition, Iassi, 27-28 mars 2009*.

Vlad, D. (à par.) « Pratiques discursives de la polémique dans le débat politique médiatique », in *Actes du colloque « Le français parlé dans les médias – les médias et le politique », Lausanne, 1-4 septembre 2009*.

Vlad, D. (à par.) « Stratégies polyphoniques à l'œuvre dans le dialogue polémique », in *Actes du XII^e Congrès International sur l'Analyse du Dialogue, Universitat Pompeu Fabra, Barcelona, 15-18 septembre 2009*.

Wolf-Mandroux, Fr. (2008) « Chassé-croisé : intégrer mobilité et pratique à l'étranger dans la formation des professeurs des écoles. Exemples d'un parcours européen et international à l'IUFM d'Orléans-Tours », Colloque international « Les dimensions européennes et internationales dans la formation des enseignants. _Regards sur les mobilités et les pratiques enseignantes », IUFM de Rennes. Vincent Marie et Nicolas Lucas (dir.), *Regards sans frontières sur la formation des enseignants*, Collection Enseigner Autrement, Editions Le Manuscrit : XX-XX.

Wolf-Mandroux, Fr. (2009) « Portfolio des compétences européennes de l'enseignant », Martine Méheut et Françoise Wolf-Mandroux, colloque de l'AEDE-France, « Comment accompagner les élèves dans leur choix d'un projet professionnel dans l'espace européen ? », publication AEDE France (Association Européenne des Enseignants), Actes du séminaire des 24-25 octobre 2008, Paris, Editions AEDE et Fondation Hypocrène, 2009.

URL : <http://www.iufm.education.fr/reseau-iufm/colloque/formation-enseignants-europe/interventions.html>

Wolf-Mandroux, Fr. (à par.) « Harmoniser à l'échelle européenne la formation des enseignants en langues grâce au PEPELF, le portfolio européen pour les enseignants en langue en formation initiale », conférence de la CDIUFM à Paris, 2008, « Enseigner en Europe. La formation des enseignants en Europe : approche comparative », Editions AEDE et Fondation Hypocrène, 2010, publication en cours.

COM

Communications orales sans actes dans un congrès international ou national

Abasq, V., Fournier, P., Fournier, J.-M. & Martin M. (2009) « A Dictionary Database of Contemporary English », PAC Workshop 2009 - *Models, Variation & Phonological Corpora*, Aix (AFF)

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

Abasq, V. (2009) « Prefix sub- and Stress Placement: sub-rules? », Colloque EPIP, Université de Savoie, Chambéry 3-5 juin 2009.

Abasq, V. (2010) « De l'utilisation des dictionnaires de prononciation pour établir un corpus », XV^e Colloque d'Anglais Oral de l'ALOES, Université Paris XIII- Villetaneuse, 26-27 mars 2010.

Abouda, L. & Perrot, M.-E. (2006) « Une question "embarrassante" en situation d'interview : modalisation et stratégies de légitimation », 3^e rencontre fribourgeoise de la linguistique sur corpus appliquée aux langues romanes, Fribourg en Brisgau (Allemagne), 14-17 septembre 2006.

Abouda, L. (2009) « Le temps des omelettes. Une nouvelle valeur pour le présent ? », 5^e Rencontres de Sémantique et Pragmatique, Université de Gabès (Tunisie), 22-24 avril 2009.

Baude, O., Jacobson, M., Tchobanov, A. & Walter, R. (2006) « Interoperability of audio corpora: the case of the French corpora », LREC 2006, Genova, Italy.

Baude, O. & Eshkol, I. (2006) « Constitution et exploitation d'un grand corpus de «données situées» *Problèmes et solutions pour les Enquêtes Socio-Linguistiques à Orléans (1968-2008)* », Colloque international 3^e rencontre Fribourgeoise de la linguistique sur corpus appliquée aux langues romanes. Fribourg Allemagne.

Baude, O. (2008) «Un grand corpus de référence du français parlé : état des lieux et perspectives», Colloque international AFLS, *Les voix du français*, Oxford, 3-5 septembre.

Baude, O. & Hriba, L. (2008) « De la variation à la norme, effets de codage dans les ESLOs », colloque international CATCOD 2008, Université d'Orléans.

Baude, O. (2009) « Les Eslos, un corpus variationniste représentatif d'une *communauté d'auditeurs* ? », Colloque international du CERLICO, L'exemple et le corpus, quel statut ? 5 & 6 juin 2009, Poitiers.

Baude, O. & Perrot, M.-E. (2009) « Les Enquêtes Sociolinguistiques à Orléans (1970-2009) : l'entretien en questions », colloque international *Pour une épistémologie de la sociolinguistique*, Montpellier, 10, 11 et 12 décembre 2009, Université Paul-Valéry.

Baude, O. & Hriba, L. (2009) « Sociolinguistique et transcription », colloque international *Pour une épistémologie de la sociolinguistique*, Montpellier, 10, 11 et 12 décembre 2009, Université Paul-Valéry.

Cavalheiro, M. (2008) « Le français véhiculaire de Ouagadougou, Burkina Faso : entre oralité et pratiques scripturales », Eurosla 18 « 18^e Conférence Internationale de l'Association Européenne

pour les langues étrangères, Atelier des doctorants, Université de Provence 1 (Aix-en-Provence), 10-13 septembre 2008.

Cavalheiro, M. (2009) « Contribution à l'élaboration d'un corpus de français parlé à Ouagadougou (Burkina Faso) », Colloque International « Pour une épistémologie de la sociolinguistique ». Université Paul Valéry, Montpellier 3, du 10 au 12 décembre 2009

Cavalheiro, M. (2010) « Usages déclaratifs et pratiques linguistiques en situation de contact de langues : l'exemple d'écoliers ouagalais du Burkina Faso », Journées des Rencontres Jeunes Chercheurs de l'ED 268 « Langage et langues », Université Paris III, 3-4 juin 2010.

Cavalheiro, M. (2010) « Le français du Burkina Faso : norme scolaire et variations. L'exemple du discours rapporté dans des écrits d'écoliers ouagalais », XXVI^e Congrès International de Linguistique et Philologie Romanes, Université de Valence (Espagne), 6-11 septembre 2010.

Chaboche, J. & **Fournié-Chaboche, S.** (2006) « Événement sportif, marketing territorial et développement local. Approche théorique et analyse sémiotique d'une publicité du Conseil Général de Vendée », Journée d'études La métalinguistique de l'événement sportif, Université d'Orléans, 2 mai 2006.

Chevrot, J.-P., **Dugua, C.** & Spinelli, E. (2007) « Acquisition des liaisons et construction morphosyntaxique », Journées PFC – Regards croisés sur la phonologie du français contemporain, Paris, 6-8 décembre 2007.

Chevrot, J.-P., **Dugua, C.** & Spinelli, E. (2008) « Acquisition of liaison and sensitivity to plural/singular orientation of nouns: a test case for a central hypothesis of the usage-based theory of language development », Boston University Conference on Language Development 33, Boston, 31 octobre-2 novembre 2008.

Cristinoi, A. (2009) « Prédire les problèmes de traduction automatique du genre dès le niveau du dictionnaire. Illustration sur le français, l'anglais et le roumain », 8^{es} Journées scientifiques du Réseau LTI, Passeurs de mots, passeurs d'espoir : lexicologie, terminologie et traduction face au défi de la diversité, Lisbonne, 15-17 octobre 2009.

Delplanque, A. (2010) « Les consonnes finales en dagara : propriétés des Éléments Phonologiques », communication présentée au colloque du Réseau Français de Phonologie, Orléans, 1-3 juillet 2010.

Dugua, C. (2006) « “un nouns”/“des nouns” ou le rôle de la fréquence sur l'acquisition de la liaison en français », Colloque international des étudiants-chercheurs en didactique des langues et en linguistique, Grenoble, 4-7 juillet 2006.

Dugua, C. & Chevrot, J.-P. (2007) « Interactions développementales entre phonologie, lexique et syntaxe autour de la liaison : un scénario d'acquisition basé sur l'usage », Colloque de l'Association Française de Linguistique Cognitive, Lille, 10-12 mai 2007.

Dugua, C. (2008) « Acquisition of liaison between determiner and noun: a usage-based model », XI International Congress for the Study of Child Language (IASCL), Edinburgh, 28 juillet - 1^{er} août 2008.

Dugua, C., Chevrot, J.-P. & Fayol, M. (2009) « Quelques aspects du développement de la liaison dans le langage des enfants de 2 à 6 ans », Présentation au sein du symposium « La liaison à l'interface entre l'oral et l'écrit », Congrès de la Société Française de Psychologie, Toulouse, 17-19 juin 2009.

Dugua, C. & Baclesse, M. (2010). « Incidence d'effets de fréquence sur l'usage de la liaison en lecture à haute voix chez de enfants de CE2-CM1 », Colloque PhonLex, Liaison et autres phénomènes de sandhi, Toulouse, 8-10 septembre 2010.

Eshkol I., Gazeau M. A. & Maurel D. (2008) « Balisage XML des entités nommées et dénommantes du corpus Eslo », First International Workshop on Cataloguing and Encoding of Spoken Language Data (CatCod 2008), Orléans, 4-5 décembre.

Eshkol I. (2008) « Enseigner le TAL à l'Université d'Orléans », Journée ATALA : « Enseignement du TAL », ENST, 26 janvier.

Fournié-Chaboche, S. (à par.) « La mise en abyme de l'écrivain : analyse d'une forme originale d'autofiction dans l'œuvre de Javier Tomeo », journée d'études « Le jeu du moi l'émoi du jeu », Université d'Orléans, avril 2008.

Fournier J.-M. (2010), « Accentuation lexicale et poids syllabique en anglais : l'analyse erronée de Chomsky et Halle », communication au Colloque annuel du Réseau Français de Phonologie, Université d'Orléans.

Garric N. & Maurel-Indart H. (2009) « Autour de la formalisation du style », Journées d'étude « La modélisation du style », Université de Tours, 10 & 11 décembre 2009.

Garric N. & Mariscal V. (2009) « L'écologie politique dans les livres-programme électoraux », Colloque international « Les discours écologistes », 26 septembre 2009, Paris-Sorbonne.

Garric N. (2008) « La médiation : un espace réflexif sur la discursivité et la construction du sens », Colloque international « Représentation du sens linguistique IV », 28-30 mai 2008, Université de Helsinki.

Gatelais, S. (2010) « Les constructions détachées en *with* et en *what with* », Workshop : Verbes et complexités verbales, 31 mai et 1^{er} juin 2010, SESYLIA, Université Paris III – Sorbonne Nouvelle.

Grenand, F. & Grenand, P. (2007). « L'humain ou l'animal ? Le cas guyanais : quand la protection de la nature peut s'exercer aux dépens des habitants humains », 4^e table ronde de l'exposition « *Bêtes et Hommes* », Grande Halle de La Villette, Paris, 24 novembre 2007.

Grenand, F. (2008) « *L'eau : usages et représentations chez les Amérindiens de l'intérieur de la Guyane* », conférence à l'usage, et sur leur demande, des professionnels de la santé en Guyane (DSDS et Croix Rouge), Cayenne, 10 p.

Grenand, F. co-organisatrice (2008) « Oyapock, un fleuve en partage, Observatoire Hommes-Millieux », communication au Colloque international multithématique, Les observatoires de l'environnement amazonien, pour le développement durable, Cayenne, 3-5 novembre 2008.

Grenand, F. (2008) « Observatório sociedades / meio ambiente: O Oiapoque, um rio compartilhado », communication dans le cadre de la 1^o Encontro dos Povos Indígenas do Amapá, norte do Pará, Suriname e Guiana Francesa, Instituto de Pesquisa e Formação em Educação Indígena, Macapá, Amapá, Brésil, 10-14 novembre 2008.

Guérin, E. & Sohmani, B. (2006) « Autour du film *L'esquive*, d'A. Kechiche, mise en lumière d'une forme d'argot contemporain », colloque *Le français une langue qui bouge*, Sceaux, 17 novembre 2006.

Hamma, B. (2006) « Sur l'invariant sémantique de la préposition *par* », Colloque International « *La préposition comme valeur* », les 4-7 avril 2006, à Villa Finaly, Florence/Italie.

Hamma, B. (2007) « Approche polyphonique des énoncés en *par* », colloque *La notion de l'altérité*, le 26 mai 2007, à l'Université de Paris 7 – Denis Diderot.

Hamma, B. & Haillet, P.-P. (2007) « *Par habitude, par exemple, par chance, etc.* : constructions de type *par X* et la notion de *contraste* », colloque *Locutions prépositives I : inventaire, classement, traduction*, le 27 janvier à Aix-en-Provence.

Hamma, B. (2007) « Etude intralinguistique des emplois de la locution *sous la coupe de* et de ses concurrentes en français », colloque *Locutions prépositionnelles, traduction et traduction automatique*, du 27 au 28 octobre 2007, à Aix-en-Provence.

Hamma, B. & Abouda, L. (2008), « De la visibilité syntaxique des modes, de l'invisibilité syntaxique des temps », *VIII^e Congrès international de linguistique française*, le 25-27 septembre 2008 Université d'Oviedo/Espagne.

Heiden, S, Jacobson, M. & **Schang, E.** (2006) « Proposition d'un mode de fonctionnement pour l'établissement de recommandations de codage et de catalogage au sein de la communauté de l'oral » Communication présentée à la journée Corpus, EPML 50, Corpus d'interaction langagière. Paris.

Jacobson, M. & **Schang, E.** (2008) « Propositions pour le catalogage des données sur les langues en contact ». Premières rencontres internationales pour le *Catalogage et le Codage de corpus oraux*, CatCod-08 Orléans, 4-5 décembre 2008.

Kihm, A. & **Rougé, J.-L.** (2009): « Sur la langue de preto et l'origine des créoles portugais d'Afrique » Communication au Groupe de recherche sur les Grammaires Créoles Université Paris 10 (30 novembre 2009)

Manuélian, H. & **Schang, E.** (2007) « L'informatisation d'un dictionnaire étymologique », Journée DECA, 12 oct 2007. Cergy Pontoise.

Maurel D., Friburger N., **Eshkol I.** & Antoine J.-Y. (2009) « Explorer des corpus à l'aide de CasSys : application au corpus d'Orléans, LingCorp'2009, Lorient, 10-12 septembre.

Moukrim, S. (2008), « Les marqueurs aspectuo-temporels du 'présent' en français, berbère (tamazight) et arabe (marocain) parlés à Orléans », Journée d'étude "Temporalité", Université de Rouen, 18 mars 2008

Moukrim, S. (2009), « *Identification de la forme verbale du présent en arabe marocain, berbère tamazight et français, parlés à Orléans* », Journée d'étude sur la marque de temps, Tours, 16 juin 2009.

Moukrim, S. (2009), « *Le projet DictAm : Dictionnaire électronique des verbes berbère-français* », 9^e Journées de L'École Doctorale SHS, Orléans, 11 et 12 juin.

Nemo, F. (2006) « Contre la modularité », Quatrièmes Rencontres de Sémantique et Pragmatique (la place du sens en linguistique). RSP4. 12 juin 2006. Orléans.

Nemo, F. (2006) *Fundamentals of Attentional Linguistics. Toward a Unified Pragmatic Metalanguage*. Paris ENST. Language, Culture and Mind Conference. LCM2. Thème Integrating Perspectives and Methodologies in Study of Language. 17 juillet 2006

Nemo, F. (2008) « Pour une sémantique non combinatoire en morphologie (et syntaxe) : introduction aux notions de *pool* et d'*intégration sémantique* », Colloque « Grammaire, morphologie, sémantique, même combat ? ». 25 septembre 2008.

Nemo, F. (2009) « Sémantique des constructions morphologiques : les routines interprétatives entre instructions et constructions ». Colloque AFFLICO Nanterre 28 mai 2009

Nemo, F. & Petit, M. (2009) « Contextes-types et formes linguistiques : les morphèmes en contexte » Colloque AFLS 2009. 5 septembre 2009 – Neuchâtel.

Nemo, F. & Petit, M. (2009) « De la prosodie en discours à la prosodie en langue, la forme prosodique entre contextes-types et emplois-types. » 10 septembre 2009. 3^e Conférence « Interfaces Discours Prosodie » (IDP '09). Université de Chicago à Paris.

Osu, S.N. (2006) « *La théorie des opérations énonciatives et la description des langues africaines* », 5^e Congrès mondial de la linguistique africaine, Addis-Abéba, 7-11 août 2006.

Perrot, M.-E. (2008), « Ecrire le chiac: choix graphiques, typographiques et lexicaux dans *Acadieman Diary* et *Acadieman Comics*, colloque international *De la France au Québec: l'écriture dans tous ses états*, Poitiers, novembre 2008.

Perrot, M.-E. (2008) « Enjeux sociolinguistiques du phénomène Acadieman », colloque international *Modes langagières dans l'histoire. Processus mimétiques et changements linguistiques*, Montpellier, juin 2008.

Perrot, M.-E. (2009) « Dix ans après : quelques observations sur l'évolution sociolinguistique du chiac », colloque international de l'AFLS, *Langue française en contexte*, Neuchâtel, septembre 2009.

Perrot, M.-E. (2010) « La dynamique du français vernaculaire des jeunes de Moncton reflétée dans deux corpus oraux (1991-2000) », II^e Journées d'Etude Internationales *Les dynamiques des français reflétées dans les corpus*, Université Paris-Nanterre.

Perrot, M.-E. (2010), « Le trajet linguistique des emprunts dans le chiac de Moncton : quelques exemples », 3^e colloque *Les français d'ici*, dans le cadre du colloque de l'ACFAS, Montréal, mai 2010.

Petit, M. (2008) « Discrimination prosodique et représentation du lexique : application aux emplois des connecteurs », 8^e Journées de l'Ecole Doctorale SHS, Université d'Orléans, 12-13 juin 2008.

Petit, M. (2009) « La diversité des fonctions d'un marqueur de structuration de discours : le cas de *enfin* en français », colloque international sur les Approches Linguistiques et Psycholinguistiques de la Structuration des Textes, ENS, Paris, 21-23 septembre 2009.

Rougé, J.-L. (2006) « L'influence lexicale portugaise sur les langues d'Afrique », Colloque ACBLPE Université de Coimbra (Portugal)

Rougé, J.-L. (2007) « Parler kriol à Ziguinchor en 2007 » Colloque SPCL/ACBLPE Université d'Amsterdam (Hollande)

Rougé, J.-L. & A. Kihm (2008) « origens da marcação da anterioridade –*ba(n)* no kriol de Guiné e Casamança » Colloque de l'ACBLPE Universidade Nova de Lisboa (Portugal)

Rougé, J.-L. & **Schang, E.** (2008) « Histoire des créoles et génétique : le cas de l'angolar. » XII^e Colloque international des Etudes Créoles. Haïti, 25-29 nov. 2008.

Rougé, J.-L. & **Schang, E.** (2009) « Eléments de reconstruction pour l'origine des créoles portugais d'Afrique » Eighth Creolistics Workshop, Pidgins and creoles in a comparative perspective, Giessen, 2-4 avril 2009.

Rougé, J.-L. (2009) « Nouvelles perspectives pour l'ambigenèse » colloque SPCL/ACBLPE Université de Cologne (Allemagne)

Schang, E. (2006) « Lexicon and polycategoriality in Saotomense : a case study? ». ACBLPE Annual Meeting, Coimbra.

Schang, E. (2007) « Anaphoric use of the demonstrative article in Saotomense », *Anaphoric uses of demonstrative expressions Workshop* at the 29th annual meeting of the Deutsche Gesellschaft für Sprachwissenschaft (Dgfs), 28 février - 2 mars 2007, Siegen.

Schang, E. (2007) « Le démonstratif *se* en saotomense », SPCL/ACBLPE Joint Meeting, Amsterdam, 18-20 juin, 2007.

Schang, E. (2008) « Les SN sans déterminants en saotomense ». ACBLPE Annual Meeting, Lisbonne, 4 juillet, 2008.

Schang, E. (2009) « Reiteration in Saotomense ». Workshop *The Grammar of Reiteration*. Amsterdam, 19 décembre 2009.

Schang, E. (2009) « Nombre et SN nus en saotomense ». Journée d'étude *Langues avec et sans articles* 2009. Paris, 6 mars 2009.
<http://www.umr7023.cnrs.fr/Programme-Program,712.html>

Schang, E. & **Rougé J.-L.** (2009) « Eléments de reconstruction pour l'origine des créoles portugais. » Communication aux *Journées d'Etude du Groupe de Recherche sur les Grammaires Créoles*, 30 nov. 2009, Univ. Paris-8.
<http://www.umr7023.cnrs.fr/-GRGC-Groupe-de-recherche-sur-les-creoles.html>

Serpollet, N. & **Poudat, C.** (2006) « Exploring the generic structure of scientific articles in a contrastive and corpus-based perspective », intervention au colloque « Digital Humanities 2006 – Premier congrès international de l'Alliance of Digital Humanities Organisations (ADHO) », Paris – Sorbonne, 5-9 Juillet 2006.

Serpollet, N. (2006) « “ Comment est-ce qu'on fait une omelette chez vous ? ” Etude des variations pragmatiques et morpho-syntaxiques du corpus oral ESLO », intervention au colloque « AFLS 2006, Variations, variétés », Association for French Language Studies, Bristol, 5-7 septembre 2006.

Serpollet, N. (2008) « “From ESLO 1 to ESLO2 : Measuring the sociolinguistic variations within two corpora of spoken French ESLO », intervention au colloque « BAAL 2008, Taking the Measure of Applied Linguistics », 41st Annual Meeting of the British Association for Applied Linguistics, Swansea 11-13 septembre 2008.

Serpollet, N. (2010) « “ Comment je fais une omelette... ? Je casse mes œufs, je les bats puis je la verse dans la poêle ” Variation des déterminants et pronoms anaphoriques dans le corpus oral et sociolinguistique ESLO », intervention au colloque « AFLS 2010, Le français : langue une, langue plurielle », Association for French Language Studies, Cambridge, Grande Bretagne, 1-3 septembre 2010.

Tellier, S., **Eshkol, I.**, Saamer, T. & Billot, S. (2010) « Morpho-syntactic labelling of an oral corpus by decomposing labels », Journée ATALA : CRF pour le TAL, Telecom ParisTech, 26 mars.

Toupin, F. (2007) « Aspects de l'influence de Donat & Priscien sur la *Grammaire* d'Ælfric », Journée d'étude « Modèles étrangers, asymétrie culturelle : les grammaires de modèles antiques », Lille 3, novembre 2007.

Toupin, F. (2009) « L'invariant à l'épreuve de la diachronie » Colloque international « Transmission et réinterprétation du passé : continuité et discontinuité », organisé par le groupe CORPUS Rouen-Amiens, mars 2009.

Vlad, D. (2007), « Le discours argumentatif polémique – un genre polyphonique par excellence », 10th International Pragmatics Conference, Goteborg, 8-13 juillet 2007.

Wolf-Mandroux, Fr. (2006) « La culture au cœur des apprentissages des langues : formation et expériences des professeurs des écoles en France », Colloque sur l'interculturel, Grenade.

AFF

Communications par affiche dans un congrès international ou national

Abasq, V. & Fournier, J.-M. (2009) «The phonology of English : Exceptions and conflict », affiche présentée à la session poster de The Sixth Old World Conference in Phonology (OCP6), University of Edinburgh, 21-24 janvier 2009.

Abasq, V., Fournier, J.-M., Fournier, P. & Martin, M. (2009) « A Dictionary Database of Contemporary English », affiche présentée à la session poster de PAC Workshop 2009 : Models, Variation & Phonological Corpora, Université d'Aix-en-Provence, 10-11 septembre 2009.

Cavalheiro, M., « Le français parlé à Ouagadougou (Burkina Faso) est-il en passe de devenir un véhiculaire urbain ? », session poster des journées de l'École Doctorale SHS, Université d'Orléans, 11-12 juin 2009.

Kanaan L., « Restatement in conversations : a strategic competence », affiche présentée à la session poster de la « Third International Conference on Spoken Communication », Naples, 23-25 février 2009.

Petit, M. (2009) « Discrimination prosodique : entre sémantique et pragmatique », affiche présentée à la session poster du colloque des doctorants et des jeunes chercheurs Coldoc 09 « L'ambiguïté dans les sciences du langage », Université Paris Ouest Nanterre la Défense, 25-26 juin 2009.

Wolf-Mandroux, Fr. (2009) « Common European Framework workshop. Measuring language learning », 6 juillet 2009, Oneida Language and Cultural Centre, Premières Nations, Ontario, Canada.

OS

Ouvrages scientifiques (ou chapitres de ces ouvrages)

Bergounioux, G. & Lamberterie (de) Ch. eds (2006) *Antoine Meillet aujourd'hui*, Collection de la Société de Linguistique de Paris, Leuven-Paris, Peeters.

Bergounioux, G. « Entre épistémologie de la grammaire comparée et figure de l'intellectuel : la situation d'A. Meillet » in Bergounioux & Lamberterie (2006) : 109-135.

Bergounioux, G. (à par.) « Affordance : de la structure de la langue à la fonction du discours », *Relire Benveniste*, R. Mahrer éd., coll. « Sciences du langage : points de vue et carrefours », Louvain, Academia-Bruylant.

Bernabé, J. & **Grenand, F.** (2006) « The Francophone Caribbean », in Ammon, U. & al., *Sociolinguistics*, An International Handbook of the Science of Language and Society, 2nd edition, Berlin, New York, Walter de Gruyter, vol. 3 : 2096-2106.

Boudreau, A. & **Perrot, M.-E.** (2010) « “Le chiac c'est du français”, représentations du mélange français/anglais en situation de contact inégalitaire », *Parlures hybrides : Regards sociolinguistiques sur quelques "bâtards linguistiques"*, Boyer, H. (ed.), Paris, L'Harmattan.

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.

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

Cadiot, P., Lebas, F. & Visetti Y.-M. (2006) « The Semantics of the Motion Verbs », *Space in Linguistics and Cognitive Categories*, Hickmann, M. & Robert, S. (eds.), Amsterdam : John Benjamins : 175-206.

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.

Cadiot, P. & Visetti Y.-M. (2009a), « Le sens commun dans la facture du proverbe » *Normativité et sens commun*, Gautier, C. & S. Laugier, S. (eds.) P.U.F. : 129-159.

Cadiot, P. (2010) *Paroles de vin, esquisses oenophiles*, Perros-Guirec, Anagrammes.

Chaboche, J., **Fournié-Chaboche S.** & Schoeny A. (2008) « Contribution à l'analyse géopolitique des Jeux Olympiques d'été : le territoire olympique à travers images et slogans officiels », in BOLI C., *Les Jeux Olympiques : fierté nationale et enjeu mondial*, Musée National du Sport Français, éd. Atlantica : 241-253.

Col, G. (2008) « Rôles de *until* et de *by* dans la mise en scène verbale », in *Préfixation, préposition, postposition. Etude de cas*, M. Paillard, M. (ed), Rennes, Presses Universitaires.

Col, G. (2010) « Des relations fonctionnelles des unités linguistiques aux relations instructionnelles », in *Espace, préposition, cognition - Hommage à Claude Vandeloise*, G. Col et C. Collin (eds). Numéro spécial de la revue *CORELA*.

URL : <http://corela.edel.univ-poitiers.fr/document.php?id=2627>

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.

Gallèpe, Th. (2006) « Pour une grammaire temporelle des didascalies. Text und Sinn - Studien zur Textsyntax und Deixis im Deutschen und Französischen. Festschrift für Marcel Vuillaume zum 60. Geburtstag », Marillier, J.-F., Behr, I. & M. Dalmas, M. (eds), Tübingen: Stauffenburg : 95 - 118.

Gallèpe, Th. (2007) « Le statut du texte didascalique : les jeux de l'entre-deux. Le texte didascalique à l'épreuve de la lecture et de la représentation », Calas F., Hamzaoui, S., El Ouri, R. & Saalaoui, T. (eds) Pessac - Tunis : Presses Universitaires de Bordeaux - Sud Editions : 23 - 38.

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.

Gallèpe, Th. (2009) « Didascalies et énonciation : "Je" interdits? », *La lettre et la scène : linguistique du texte de théâtre*, C. Despierres, H. Bismuth, M. Krazem & C. Narjoux (eds), Dijon: EUD – Langages, 43-57.

Garric N., Léglise I., Canut E. & Desmet I. (2006) « Introduction », *Applications et Implications en Sciences du langage*, Paris, L'Harmattan : 9-15.

Garric N. & Léglise I. (2006) « Le discours patronal, discours de propagande », *Aspects linguistiques du texte de propagande*, Paris, L'Harmattan : 133-136.

Garric N. (2010) « L'identité événementielle : "raz de marée", "tsunami", "shoah" », *Construction d'identité et processus d'identification*, Peter Lang : 275-294.

Grenand, F., Bahuchet, S. & Grenand, P. (2006) « Environment and peoples in French Guiana : ambiguities in applying the laws of the French Republic », in Roué, M. « Cultural Diversity and Biodiversity », *International Social Science Journal* (58)187 : 49-58.

Grenand, F. (2008) « Nommer son univers : Pourquoi ? Comment ? Exemples parmi des sociétés amazoniennes », in Prat, D., Raynal-Roques, A. & Roguenant, A., *Peut-on classer le vivant ? Linné et la systématique aujourd'hui*, Paris, Belin : 119-130.

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.

Grenand, P., Tilkin-Gallois, D. & **Grenand, F.** (2009) « Un peuple Tupi-Guarani, les Wayâpi », in Grenand, F. (éd. en chef) *Encyclopédies palikur, wayana, wayâpi : langue, milieu et histoire*, fascicule 0 : 86-115, Orléans-Paris, PUO-CTHS, Paris.

Grenand, F. (2010) « Nomear seu Universo: Porque? Como? Alguns exemplos de sociedades amazônicas », trad. J. Cabral de Oliveira, *Cadernos de campo* 18, São Paulo, USP.

Grenand, F. associate editor (sans date). *South American Indian Languages*, Intercontinental Dictionary Series, Bernard Comrie (chief ed.), Irvine University.
URL : <http://lingweb.eva.mpg.de/ids/>.

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

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 (224 pages). 143-158.

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 (506 pages) : 415-448.

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.

Nemo, F. (2007) « La sémantique peut-elle être catégorielle ? », *La représentation du sens en linguistique 2*, Bouchard, D., Evrard, I. & Vocaj E. eds, Bruxelles, Boeck-Duculot : 35-52.

Petit, M. (à par.) « Prosodie et sémantique des mots : pour une morphologie élargie », in *Liens linguistiques. Études sur la combinatoire des composants*, coll « Sciences pour la communication », Lang.

Rougé J.-L. (2006) « L'influence Manding sur la formation des créoles du Cap-Vert et de Guinée Bissau et Casamance » *Cabo Verde : Origens da sua sociedade e do seu crioulo* Lang, Holm, Rougé & Soares eds, Tubingen, Gunter Narr Verlag : 63-74.

Rougé J.-L. (2008) « A inexistência de crioulo no Brasil », *Africa no Brasil*, J.-L. Fiorin & M. Petter eds, São Paulo, Editora contexto : 63-74.

Schang, E. (à par.) « Reiteration in Saotomense : Issues of the syntax-semantics interface », *The Grammar of Reiteration*, Aboh, E.O., Smith, N. & Zribi-Hertz, A. eds, Amsterdam, J. Benjamins.

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.

OV

Ouvrages de vulgarisation (ou chapitres de ces ouvrages)

Bergounioux, G. (à par.) « *Sémitisme* : d'un concept linguistique à un argument raciste », A. Meddeb et B. Stora eds, *Histoire des relations entre juifs et musulmans du Coran à nos jours*, Paris, A. Michel.

Bourdier, Ph. & Caglar, P. (2006) *Contes de Charles Perrault illustrés par Gustave Doré : langage verbal et images*, col. « Réseau diagonales », éd. Ellipses, Paris, 192 p.

Bourdier, Ph. (2006) *La fête : BTS : variations picturales sur le motif biblique de l'Adoration du veau d'or, Jan Steen et Nicolas Poussin*, Calais, E éd., éd. Ellipses, Paris : 103-4 ; 109-10.

Garric, N. & Calas F. (2007) *Introduction à la Pragmatique*, Paris, Hachette, « Linguistique ».

Garric, N. (2007) *Introduction à la linguistique*, Paris, Hachette, « Linguistique ».

Grenand, F. (2007) « Des Jaguars, des Grenouilles et des Hommes, ou Comment trouver sa place dans l'univers : un exemple amazonien », in DESPRET, V., *Bêtes et Hommes : catalogue de l'exposition à la Grand Halle de La Villette*, sept. 2007-janv. 2008, Paris, Gallimard : 16-17.

Grenand, F. (2006). « Savoir boire et savoir vivre chez les Wayãpi de Guyane et du Brésil », in Erikson, Ph. *La pirogue ivre : bières traditionnelles en Amazonie*, Nancy, Musée de la Brasserie : 13-23 ; 145-146.

Grenand, F. (2009). « La langue wayãpi », *Les langues de Guyane*, Renault-Lescure, O. & Goury, L. La Roque d'Anthéron, Ici et Ailleurs/Vents d'Ailleurs : 100-109.

Iglesias Ovejero, A. (2006) *Blasones populares del antiguo partido de Ciudad Rodrigo : nombres, sobrenombres, tópicos y matracas pueblerinas, según el refranero, el cancionero, el romancero y las leyendas de la crónica oral*, Centro de Estudios Mirobrigenses (CSIC), Ciudad Rodrigo 59 pp.

JOURNEES d'ETUDES, ECOLES THEMATIQUES, ATELIERS

Abouda, L. (2008) « De la distorsion temporelle », journée d'étude « Temporalité » organisée par Laurent Gosselin (Université de Rouen).

Abouda, L. & Moukrim, S. (2010) « Deux groupes de formes temporelles en français », journée « Temporalité » organisée par S. Ossu et F. Toupin (LLL), Tours, 5 mars 2010.

Baude, O. (2006) « Diffusion des corpus oraux, problèmes juridiques et déontologiques », Ecole thématique du CNRS ELCO, Nantes.

Baude, O. (2007) « Mutualiser des corpus oraux, aspects juridiques et déontologiques » Journées d'étude CORPAFROAS, Corpus Oral en langues Afroasiatiques : Analyse Prosodique et Morphosyntaxique, 15 février, Paris.

Baude, O. (2007) « Constituer et exploiter un corpus d'interactions- aspects juridiques et éthiques - », Ecole thématique du CNRS CONTACI, Lyon.

Baude, O. (2007) « Le corpus d'Orléans », Journée d'études de la BnF, *Autour du Français Parlé : de Brunot à nos jours. De l'archivage à l'exploitation.*

Baude, O. (2008) « Les enquêtes sociolinguistiques à Orléans, Base et corpus ». Ecole thématique CNRS I_DOCORA Interaction : DONnées, CORpus, Analyse 23 au 27 juin 2008, Fourvière – Lyon.

Baude, O. (2008) « Outiller la sociolinguistique : une démarche réflexive autour du corpus des ESLOs (1968-2008) », Ecole thématique du CNRS / TGE ADONIS *Préservation et diffusion numériques des sources de la recherche en sciences humaines et sociales* 19-24 octobre 2008, Fréjus (Var)

Baude, O. (2008) « Valorisation de corpus oraux : du terrain au portail », Ecole thématique du CNRS / tge ADONIS *Préservation et diffusion numériques des sources de la recherche en sciences humaines et sociales* 19-24 octobre 2008, Fréjus (Var).

Baude, O. & Hriba, L. (2008) « *Les Enquêtes SocioLinguistiques à Orléans (1968-2008)* : choix méthodologiques pour un corpus prototypique », Journée d'étude *Parole*, FORELL, Poitiers.

Bourdier, Ph. (2009) « Les interactions lecture-écriture et le cas du récit au cycle 3 de l'école élémentaire », Journée d'étude consacrée à l'écriture à l'école, CRDP Montpellier-Rectorat Montpellier, Montpellier, 26 mars 2009.

Nemo, F. (2007) « Stabilité et instabilité en sémantique linguistique », Université de Limoges, conférence invité au séminaire de linguistique, 8 mars 2007

Nemo, F. (2008) « Enoncés, contributions, discours », conférence doctorale à l'Université de Cluj Napoca. 27 juin 2008

Nemo, F. (2009) « Ressorts et supports de "l'argumentativité linguistique" ». Journée "Les sémantiques argumentatives aujourd'hui". Université de Nantes. 7 février 2009.

Nemo, F. (2009) « Routines interprétatives, constructions grammaticales et constructions discursives ». Université de Nova Lisboa. Juillet 2009.

Nemo, F. (2010) « Dualité du lexique, bases morphémiques et adressage lexical », Colloque « Le thème perceptif en linguistique », EHESS. 12 mars 2010

Osu, S.N. (2008) « *Diversité et universalité des langues : y a-t-il des langues francophones ?* » Table ronde « Journée des éditeurs » à l'Université d'Angers.

Rougé, J.-L. (2007) « Les créoles portugais : états des lieux ». Communication au Groupe de Recherche sur les Grammaires Créoles Université Paris 10 (14 mai 2007).

Rougé J.-L. & Schang, E. (2009) « Eléments de reconstruction pour l'origine des créoles portugais. » Communication au Groupe de recherche sur les Grammaires Créoles Université Paris 10 (30 novembre 2009)

DO

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

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'EHESS.

Bourdier, Ph. dir. (2010) « L'enseignement du cinéma dans les années 50 en France : l'exemple des actions de *Film et jeunesse* », *Images à l'école, image de l'école, 1880-1960*, éd. Musée de l'école, juin 2010 : 35-45.

Col, G. et Collin, C. (2010). *Espace, préposition, cognition - Hommage à Claude Vandeloise*. Numéro spécial de la revue *CORELA*. (à par.)

Capiberibe A., **Cristinoi A.** & Grenand P. dir. (2009) *Encyclopédie Palikur - Fascicule 1*, Paris, Éd. du CTHS, Orléans / P.U.O.

Delesse C., Lowrey, B. & **Toupin, F.** (2010) *Actes du premier Colloque Bisannuel de Diachronie de l'Anglais*. Paris, Publications de Association des Médiévistes Anglicistes de l'Enseignement Supérieur 31 (251 pages).

Delplanque, A. (2006) « Les verbes d'apparence », numéro spécial de la revue *CORELA*.
URL : <http://edel.univ-poitiers.fr/corela/>

Garric N. & Leglise I., Canut E., Desmet I. (2007) *Applications et implications en sciences du langage*, Paris, L'Harmattan.

Garric N. & **Longhi J.** (2009) « L'analyse linguistique de corpus discursifs : des théories aux pratiques, des pratiques aux théories », *Les Cahiers du LRL*, n°3, Université de Clermont-Ferrand.

Garric N. & Léglise I. (à par.) *Les discours d'experts*, Peter Lang.

Grenand, F. & Camargo, E., Capiberibe, A., **Cristinoi, A.**, Gallois, D., Grenand, P. « *Encyclopédies palikur, wayana, wayãpi : langue, milieu et histoire* », fascicule 0 de l'*Encyclopédie des peuples de Guyane*, Paris/Orléans, PUO-CTHS, 165 p.

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

Iglesias Ovejero, A. dir. (2006) *Cahiers du P.R.O.H.E.M.I.O.* 7 et 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. Presses Universitaires d'Orléans, 315 pp. et 451 pp.

Iglesias Ovejero, A., Fernandez, J., Giraud, F., Iglesias, C. 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 de las III Jornadas Internacionales de Estudio de El Rebollar (VII Colloque PROHEMIO, 20-22 juillet 2007), P.U.O., 406 pp. et 458 pp.

Iglesias Ovejero, A., Giraud, F. & Sánchez Valdelvira, B. dir. (2010) *Cahiers du PROHEMIO* 11, DVD, *Historia, habla y cultura tradicional*. Actas de las IV Jornadas Internacionales de Estudio de El Rebollar (VIII Coloquio PROHEMIO, 17-19 juillet 2009, P.U.O. 560 pp.

Lafont-Terranova, J. & Colin, D. (éds.) (2006) *Didactique de l'écrit : La construction des savoirs et le sujet-écrivain. Actes de la journée d'étude du 13 mai 2005 CORAL Université d'Orléans, IUFM d'Orléans-Tours (France)*, Namur, CEDOCEF / Presses universitaires de Namur (Diptyque 5).

Loiseau, M., Abouzaïd, M., Buson, L., Cavalla, C., Djaroun, A., **Dugua, C.**, Ghimenton, A., Goossens, V., Lebarbé, T., Nardy, A., Rinck, F. & Surcouf, C. dir. (2008). *Autour des langues et du langage – Perspective pluridisciplinaire*, Grenoble, Presses Universitaires de Grenoble.

Osu, S., Col, G., Garric, N. & Toupin, F. (2010). *Construction d'identité et processus d'identification*. Bern : Peter Lang.

Lang J., Holm J., **Rougé, J.-L.** & Soares M.-J. (200X) *Cabo Verde : Origens da sua sociedade e do seu crioulo*. Ed. Gunter Narr Verlag Tübingen

Régis, J.-P. & Toupin, F. eds (2006) *Travaux de Diachronie 3* (Actes du Deuxième Colloque de diachronie de l'anglais à Tours, mars 2001). Tours, Presses universitaires François Rabelais (105 pages).

AP

Baude, O., Alessio, M., (2008) «Les corpus de la parole, patrimoine immatériel et langues de France», in *Culture & Recherche*, n° 116 et 117, Ministère de la Culture et de la communication, Paris. p 42-43.

Baude, O. (2008) «Les “bonnes pratiques” de constitution et d'exploitation de corpus oraux, un exemple d'initiative fédérative pour une communauté spécifique», Atelier ANTHROPONET: *champ documentaire et champ scientifique : Quelles pratiques et quels standards dans l'indexation de corpus scientifiques multimédia ?*, Orléans, 26-27 juin 2008.

Baude, O. (2008) « L'oral un domaine à exploiter » conférence au salon Expolangues, 7 février, Paris.

Baude, O. (2009) « découvrir les langues de France : le site corpus de la parole » Colloque international du CERLICO, *L'exemple et le corpus, quel statut ?* 5 & 6 juin 2009, Poitiers.

Bergounioux, G. & Fryba-Reber, A.-M. (2006) édition du *Journal* d'A. Meillet in Bergounioux & Lamberterie : 21-86.

Bergounioux, G. (2006) « Variations », interview de Pierre Encrevé, in *Faire signe*, Presses Universitaires d'Orléans : 57-93

Bergounioux, G. (2006) « La description des langues : phonétique et sémantique », *Annuaire de l'EHESS* : résumé des cours et conférences : 656-657

Bergounioux, G. (2006) « Les Enquêtes Socio-Linguistiques à Orléans » (avec O. Baude), DGLFLF, Ministère de la Culture et de la Communication, *Langage et Cité* : 2-3

Bergounioux, G. (2006) « Gnôthi seuton », préface de C. Collin, *Le sujet dans l'injonction*, Presses Universitaires de Rennes : 15-17

Bergounioux, G. (2006) « L'enseignement des sciences du langage et la didactique de l'écrit » – préface à *Dyptique* 5, édité par J. Lafont-Terranova et D. Colin, Namur, Presses Universitaires de Namur : 5-6

Bergounioux, G. (2006) « Une de mes amours, un de mes travaux » in *Penser/Rêver* 10 (entretien avec M. de Azambuja, M. & M. Gribinski), Editions de l'Olivier : 263- 286

Bergounioux, G. (2006) CR de *Gaston Paris et la philologie romane* d'U. Bähler, *Bulletin de la Société de Linguistique de Paris* CI, 2 : 129-137

Bergounioux, G. (2006) CR de *Saussure : la langue, l'ordre et le désordre* d'A.-J. Pétrouff, *Bulletin de la Société de Linguistique de Paris* CI, 2 : 199-202

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. (2007) CR de *Charles Bally (1865-1947) : historicité des débats linguistiques et didactiques* par J.-L. Chiss (ed.), *Bulletin de la Société de Linguistique de Paris* CII, 2 : 5 – 7

- Bergounioux, G.** (2007) CR de la *Correspondance avec R. Jakobson et autres écrits* de N. S. Troubetzkoy, éditée par P. Sériot, *Bulletin de la Société de Linguistique de Paris* CII, 2 : 11 – 18
- Bergounioux, G.** (2007) CR d'*Allegro ma non troppo. Invitation à la linguistique* de Cl. Normand, *Bulletin de la Société de Linguistique de Paris* CII, 2 : 72-76
- Bergounioux, G.** (2007) CR de *Motifs et proverbes. Essai de sémantique proverbiale* d'Y.-M. Visetti et P. Cadiot, *Bulletin de la Société de Linguistique de Paris* CII, 2 : 162 - 164
- Bergounioux, G.** (2008) CR d'*A la recherche de Ferdinand de Saussure* par M. Arrivé, *Bulletin de la Société de Linguistique de Paris* CIII, 2 : 88 – 90
- Bergounioux, G.** (2008) CR de *Lévi-Strauss mot à mot. Essai d'idiographie linguistique* par H. Curat, *Bulletin de la Société de Linguistique de Paris* CIII, 2 : 95 – 98
- Bergounioux, G.** (2008) CR des *Cahiers Ferdinand de Saussure* 59, *Bulletin de la Société de Linguistique de Paris* CIII, 2 : 99 – 102
- Bergounioux, G.** (2008) CR du *Français au Levant, jadis et naguère. A la recherche d'une langue perdue.* par C. Aslanov, *Bulletin de la Société de Linguistique de Paris* CIII, 2 : 294 – 297
- Bergounioux, G.** (2009) « Etre étudiant en sciences du langage, pour faire quoi ? » in *Sciences du langage et demandes sociales*, Ch. Hudelot et Ch. Jacquet-Pfau eds, Limoges, Lambert-Lucas : 181-184
- Bergounioux, G.** (2009) CR de *Le linguiste et l'inconscient* par M. Arrivé, *Bulletin de la Société de Linguistique de Paris* CIV, 2 : 56 – 58
- Bergounioux, G.** (2009) CR *Du côté de chez Saussure* par M. Arrivé (ed.), *Bulletin de la Société de Linguistique de Paris* CIIV, 2 : 58 – 65
- Bergounioux, G.** (2009) CR de *La théorie du langage dans l'enseignement de Jacques Lacan* par V. Calais, *Bulletin de la Société de Linguistique de Paris* CIV, 2 : 65 – 66
- Bergounioux, G.** (2009) CR d'*Une historiographie engagée* par J.-Ph. Saint-Gérand et Y. Portebois, *Bulletin de la Société de Linguistique de Paris* CIV, 2 : 273 – 276
- Bergounioux, G.** (2009) CR de *La linguistique sociologique de Charles Bally* par Cl. Forel, *Bulletin de la Société de Linguistique de Paris* CIV, 2 : 281 – 284
- Bergounioux, G.** (2010) « La langue des Indo-Européens ? » « Les Indo-Européens » *Dossiers d'Archéologie* 338 : 14-21
- Cadiot, P.** (2008) CR de *The magic Prism* par H. Wettstein, *RSP* 23 : 143-146.
- Cristinoi, A.** (2008) CR *Le discours en interaction* de C. Kerbrat-Orecchioni (Armand Colin, 2005, 365 pages), *RSP* 23 : XX-XX
- Cristinoi, A.** (2008) CR *A Semantic Approach to English Grammar* de R.M.W. Dixon, *RSP* 23 : XX-XX.

Delplanque, A. (2009) « Typologie linguistique et histoire du peuplement : le cas des langues gur du Burkina-Faso », pdf en ligne sur HAL [00429049, v2].

URL : <http://hal.archives-ouvertes.fr/>

Eshkol, I. (2010) *Guide d'annotation morpho-syntaxique*, à intégrer dans la site ESLO du LLL.

Grenand, F. & Grenand, P. (2007). « L'humain ou l'animal ? Le cas guyanais : quand la protection de la nature peut s'exercer aux dépens des habitants humains », 4^e table ronde de l'exposition « *Bêtes et Hommes* », Grande Halle de La Villette, Paris, 24 novembre 2007.

Grenand, F. (2008) « *L'eau : usages et représentations chez les Amérindiens de l'intérieur de la Guyane* », conférence à l'usage, et sur leur demande, des professionnels de la santé en Guyane (DSDS et Croix Rouge), Cayenne, 10 p.

Grenand, F. co-organisatrice (2008) « Oyapock, un fleuve en partage, Observatoire Hommes-Millieux », communication au Colloque international multithématique, Les observatoires de l'environnement amazonien, pour le développement durable, Cayenne, 3-5 novembre 2008.

Grenand, F. (2008) « Observatório sociedades / meio ambiente: O Oiapoque, um rio compartilhado », communication dans le cadre de la 1^o Encontro dos Povos Indígenas do Amapá, norte do Pará, Suriname e Guiana Francesa, Instituto de Pesquisa e Formação em Educação Indígena, Macapá, Amapá, Brésil, 10-14 novembre 2008.

Grenand, F. & Nemo, F. (2008) « Territoire en mutation » et « Encyclopédie thématique des langues de Guyane », « Savoirs, connaissances et développement », *Triptyque 2* : 6-7.

Grenand, F. (2009) « Nommer l'univers qui nous entoure : pourquoi ? Comment ? », conférence dans le cadre des Vendredis d'Irista, réseau de diffusion scientifique et technique de la Guyane, Cayenne, 30 janvier 2009.

Grenand, F. co-organisatrice (2009) « Gestion et évolution des terres et de l'agriculture », communication dans le cadre de la 2^e Rencontre des Peuples Amérindiens transfrontaliers de l'Amapá, du Nord-Pará, du Surinam, de Guyana et de Guyane française, Iepé-CNRS, Saint-Georges de l'Oyapock, Guyane française, 2-6 décembre 2009.

Grenand, F. (2010) « Le pont sur l'Oyapock : enjeux et défis », conférence et projection du film de Dalaise, M. *Oyapock, un fleuve en partage*, 52', CNRS-images Paris, dans le cadre des Vendredis d'Irista, réseau de diffusion scientifique et technique de la Guyane, Cayenne, 20 mai 2010.

Guérin, E. (2009) « Oral/écrit : une opposition saillante qui trouble l'objet d'enseignement du FLM », Séminaire de l'équipe « Dynamiques des usages » du laboratoire MoDyCo, Université Paris X, 25 mars 2009.

Guérin, E. & Sohmani, B. (2007) « Les pratiques langagières des jeunes locuteurs en débat », *La nuit des chercheurs*, Cité Universitaire Internationale de Paris, 28 septembre 2007.

Guérin, E. (2007). « Propositions pour une didactique de la grammaire qui rendrait compte de la nécessaire variabilité de la langue », Journée d'étude *Enseignement du français, variation, plurilinguisme*, Université de Cergy Pontoise, 12 juin 2007.

Guérin, E. (2010). « Entretiens avec des enfants d'Épinay sous Sénart – une étude sociolinguistique à l'usage des acteurs de l'enseignement », compte rendu du rapport remis au GIP-RE spinolienne, *Langage et Sociétés*, 131.

Nemo, F. (2009) CR *Minimal Semantics* d'E. Borg, *RSP* 25-26 : XX-XX.

Osu, S.N. (2006) CR du *Dictionnaire igbo-français suivi d'un index français-igbo* d'Ugochukwu, F. & Okafor, P., Paris-Ibadan (Nigeria), Karthala/IFRA, *Journal des Africanistes* 76-2 : 172-175.

Osu, S.N. (2009) CR d'*Antoine Culioli : un homme dans le langage* de Ducard, D. & Normand, Cl., *Journal des Africanistes* 79-1 : 303-307.

Petit, M. et al. (2007) « La lumière nous en fait voir de toutes les couleurs », *Microscop* hors-série 16 : pp 46-47.

Petit, M. (à par.) Compte rendu de *Pragmaticalisation et marqueurs discursifs. Analyse sémantique et traitement lexicographique*, *Revue de Sémantique et Pragmatique* 25-26.

Petit, M. (à par.) « L'intonation au service de l'interprétation », *Pour la Science*.

Toupin, F. (2006) Biographie de Henri Adamczewski, *Encyclopedia of Language and Linguistics*, 2nd edition, Oxford, Elsevier.

Toupin, F. (2008) « L'Invariant a-t-il un avenir ou comment la diachronie peut aider au retour du refoulé », *Mélanges de langue, littérature et civilisation offerts à André Crépin à l'occasion de son 80^e anniversaire* (Médiévales 44), Presses du Centre d'Etudes Médiévales de l'Université d'Amiens : 303-11.

Vlad, D. (à par.), « Sur la polémique au sein du discours scientifique », in *Mélanges, Interactions Culturelles Européennes*, Université de Bourgogne, Dijon.

TH

Thèses

Poudat, C. (2006) *Etude contrastive de l'article scientifique de revue linguistique dans une perspective d'analyse des genres*, Université d'Orléans, dir. G. Bergounioux.

Abasq, V. (2007) *Préfixation et particules adverbiales en anglais contemporain : étude du comportement accentuel*, Université de Tours, dir. J.-M. Fournier.

Cloiseau, G. (2007) *Métaphores vives et patrons prosodiques : analyse contrastive dans un corpus de musiciens*, Université d'Orléans, dir. G. Bergounioux.

Cristinoi, A. (2007) *Analyse contrastive des indices morphosyntaxiques nominaux de genre et de nombre en vue d'une approche typologique de la traduction automatique. Application sur le français, l'anglais et le roumain*, Université d'Orléans, dir. G. Bergounioux.

Dugua, C. (2006). *Liaison, segmentation lexicale et schémas syntaxiques entre 2 et 6 ans. Un modèle développemental basé sur l'usage*. Thèse de doctorat, Université Stendhal Grenoble3, Grenoble.

Girard, I. (2007) *Isomorphisme en anglais contemporain. Etude de quelques suffixes*, Université de Tours, dir. J.-M. Fournier.

Talmenssour, A. (2007) *Etude de la parémiologie en arabe marocain*, Université d'Orléans, dir. P. Cadiot.

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

Feugain, A. M. (2008) *Iconologie et iconographie : analyse contrastive des affiches de propagande pendant la II^e République et la guerre civile espagnole (1931-1939)*, Université d'Orléans, dir. A. Iglesias et F. Nemo.

Vlad, D. (2008) *La polyphonie dans les structures de la langue. Le cas du discours polémique*, Université d'Orléans et Université de Cluj-Napoca, dir. L. Florea et F. Nemo

Petit, M. (2009) *Discrimination prosodique et représentation du lexique : application aux emplois des connecteurs*, Université d'Orléans, dir. F. Nemo.

Moukrim, S. (2010) *Morphosyntaxe et sémantique du « présent » : une étude contrastive à partir de corpus oraux (ESLO/LCO) arabe (marocain), berbère (tamazight) et français*. Université d'Orléans, dir. L. Abouda et G. Bergounioux

Autres formes de production (brevets, ouvrages, rapports d'expertise)

Baude, O. (2006-2010) Rédacteur en chef de *Langues & cité*, bulletin de la Délégation générale à la langue française et aux langues de France, Ministère de la Culture.

Baude, O. (2006-2010) Directeur scientifique du *Corpus de la parole*. DGLFLF-Ministère de la Culture et de la communication & CNRS. Corpus en ligne : <http://www.corpusdelap parole.culture.fr/>

Bergounioux, G. ed. (2006) *Pour un Référentiel Européen d'Enseignement de la Linguistique* (avec l'équipe pédagogique des sciences du langage à Orléans), brochure, 56 pages.

Bergounioux, G. (2010) contribution au rapport de conjoncture de la section 34 du CNRS (phonologie, histoire de la linguistique, MSH).

Bergounioux, G. (2009) Claire Le Feuvre, *Le Vieux-Slave*, collection « Langues du Monde » publiée par la Société de Linguistique de Paris, Leuven-Paris, Peeters, lecture, correction et préparation du manuscrit.

Bergounioux, G. (2010) Claire Blanche-Benveniste, *Le Français parlé*, collection « Langues du Monde » publiée par la Société de Linguistique de Paris, Leuven-Paris, Peeters, lecture, correction et préparation du manuscrit.

Bergounioux, G. (à par.) A. Montaut, *Le Hindi*, collection « Langues du Monde » publiée par la Société de Linguistique de Paris, Leuven-Paris, Peeters, lecture, correction et préparation du manuscrit.

Cavalheiro, M. (2007-2011) Base de données sur le français parlé en milieu scolaire à Ouagadougou.

Col, G. (2006-2010) Responsable de publication de la revue *CORELA – Cognition / Représentation / Langage* (<http://revue-corela.org>) : 4 volumes de 2 numéros chacun.

Cristinoi, A. trad (2010) *George la vânătoare de comori prin Cosmos*, București, Humanitas, traduction de l'anglais de *George's Cosmic Treasure Hunt* de Lucy & Stephen Hawking.

Cristinoi, A. trad (2008) *George și cheia secretă a Universului*, București, Humanitas, traduction de l'anglais de *George's Secret Key to The Universe* de Lucy & Stephen Hawking.

Cristinoi, A. trad (2006) *Animalul timpului*, București, Editura Omonia, traduction du français de *L'animal du temps* de Valère Novarina.

Maurel D., Friburger N., **Eshkol I.** (2009) Corpus étiqueté ESLO1 des entités nommées et des entités dénommantes (36,73 Mo).

Eshkol I., Tellier I., Saamer T. (2010) Sous-corpus ESLO, 50.000 mots étiquetés avec informations morpho-syntaxiques.

Fournier, J.-M. (2008), « Y a-t-il de la diachronie dans la synchronie ? », communication aux Journées d'Étude du *Laboratoire Ligérien de Linguistique* sur la temporalité.

Fournier, J.-M. (2007), « L'identification phonologique de l'étranger en anglais contemporain », communication aux Journées d'Étude du *Laboratoire Ligérien de Linguistique* sur la construction d'identité.

Fournier, J.-M. (2006), « Traitement phonologique des noms propres anglais », communication dans le cadre des travaux du groupe *Langues et Représentations* sur Construction d'identité et processus d'identification

Grenand, F. (2007) supervision de : Chapuis, J. *L'ultime fleur, Ekulunpi tihmelë...essai d'ethnosociogenèse wayana*, fascicule 1 de l'encyclopédie wayana, *Encyclopédie des peuples de Guyane*, Orléans, PUO, 164 p.

Grenand, F. (2008) supervision de : Nimuendaju, C. *Les Indiens Palikur et leurs voisins*, trad. Steiner, W. & Lecler, J. introduction et annotations Grenand, P., Fascicule 1 de l'encyclopédie palikur, *Encyclopédie des peuples de Guyane*, Paris/Orléans, PUO-CTHS, 188 pages.

Guérin, E. (2009). *Pratiques langagières et rapport à l'enseignement du français : étude auprès des enfants de la ville d'Épinay sous Sénart*, projet financé par le Groupement d'Intérêt Public-Réussite Educative spinolienne, rapport remis au GIP-RES en juin 2009.

Lafont-Terranova, J. & Colin, D. (2006) « La question de la norme dans le discours des enseignants de collège », in Lafont-Terranova, J. & Colin, D. (éds.), *Didactique de l'écrit : La construction des savoirs et le sujet-écrivain*, Namur, CEDOCEF / Presses universitaires de Namur (Diptyque 5) : 105-131.

Perrot, M.-E. (2007 sq.) Participation au projet *Idéologies et représentations linguistiques dans les textes écrits sur l'Acadie de la fin du 19^e siècle à la période contemporaine*, corpus de textes écrits répertoriés dans des bases de données et en partie numérisés) sous la direction d'Annette Boudreau.

Perrot, M.-E. (2010 sq.) Participation au projet *La construction des idéologies linguistiques en Acadie des Maritimes et en Louisiane* sous la direction d'Annette Boudreau.

Rougé J.-L. (2007) « Construction de la langue, construction de l'identité », *Actes des journées de l'ADAMIF*.

Rougé J.-L. (2008) « Langues en contact un projet pilote à Orléans » in *Langue et cité* 12.

Rougé J.-L. (2010) « Une richesse du français : le contact de langues » in *Langue et cité* 16.

Schang, E. (2009) Porteur et rédacteur du projet C02 (COréférence dans les Corpus Oraux), Universités d'Orléans et Tours.

Wolf-Mandroux, Fr. (2007) *Les compétences des enseignants de langues étrangères dans les écoles primaires et les collèges en Croatie*, ouvrage collectif traduit en 4 langues (anglais, français, croate et allemand), Publication Projet Tempus CD-18049-2003.

Wolf-Mandroux, Fr. (2007) *Langues étrangères en primaire : formation d'enseignants. Rapports et conclusions*, ouvrage collectif traduit en 4 langues (anglais, français, croate et allemand). Publication Projet Tempus CD-18049-2003.

Wolf-Mandroux, Fr. (à par.) « Reflections on the use of the EPOSTL (European Portfolio for Student Teachers of Languages): the example of a research and development action in France and three European countries (Bulgaria, Finland and Czech Republic) », bilan de la recherche-action sur le PEPELF.

ANNEXES

Statuts et règlement du laboratoire

UMR de linguistique : tableau récapitulatif

La recherche à la BnF

Proposition de convention avec la BnF : CODEX

CAT COD : Appel à colloque 2008

Revue de Sémantique et Pragmatique (2007-2010)

Revue CORELA

CERLICO : Présentation

CERLICO : 24^e colloque (Tours 2009)

Collaboration avec le LATTICE

Rapport de la thématique : « Construction du sens »

Journée d'étude sur la temporalité : programme

ESLO : VARILING (présentation)

GUYANE : Encyclopédies

GORDF : Séminaire

PROHEMIO : dépouillement des ouvrages pour un dictionnaire des figements

PROHEMIO : Sommaires de la revue

CRITERR (Orléans) : recherches en cours

RFP : Appel à communication et programme

MSH : Projet collaboratif

Statuts et règlement du LLL contrat 2008 – 2011
(adoptés : assemblée générale du 17 janvier 2008)

Laboratoire Ligérien de Linguistique, EA 3850 / ILF – CNRS

Directeur Gabriel Bergounioux (Université d'Orléans)

Directeur-adjoint, responsable de l'équipe de Tours, Jean-Michel Fournier

1. Objectifs

Recherche linguistique, formation des doctorants, support de formation de master, convergence des recherches sur le plan régional (Fédération FELINCO – Master international FELINCO), partenariats multiples selon les objectifs scientifiques.

2. Organigramme thématique

A Tours (3 axes et 1 atelier thématique)

- morphophonologie
- analyse du discours
- diachronie de l'anglais
- construction du sens

A Orléans (5 équipes)

- ESLO
- Langues d'Afrique et créoles
- Langues de Guyane
- GORDF
- PROHEMIO

3. Organismes de rattachement

Université François-Rabelais, Ecole Doctorale SHS, UFR Lettres & Langues)

Université d'Orléans, Ecole Doctorale SHS, UFR Lettres, Langues et Sciences Humaines

4. Membres : critères de définition des membres permanents et associés
Membres permanents : rattachement principal au laboratoire

Membres associés : collaborateurs sur programmes : professeurs émérites, doctorants, enseignants-chercheurs en rattachement secondaire, PRAG, enseignants de l'IUFM, chercheurs étrangers...

5. Droits et obligations des membres

- Droits : accès aux locaux et au matériel, support financier de leur activité de recherche, voix délibératives (ou consultatives pour les associés) à toutes les décisions du groupe de recherche (administratives, organisationnelles, financières)
- Obligations
 - ➔ Participation aux séminaires et implication dans des projets de recherche au nom du laboratoire
 - ➔ Encadrement doctoral : respect de la charte des thèses, des règles de l'Ecole Doctorale
 - ➔ Signature des publications sous l'appellation Laboratoire Ligérien de Linguistique

6. Institution et fonctionnement

- Assemblée Générale

sont membres de droit de l'assemblée générale tous les membres permanents du laboratoire. Les doctorants du laboratoire et les membres associés participent avec voix consultative. L'assemblée générale se réunit au moins une fois par an, élit le Conseil de laboratoire, adopte le bilan de l'année écoulée et les perspectives pour les années à venir.

- Conseil de laboratoire

Le Conseil de laboratoire est composé de 10 membres, en règle générale les membres des bureaux de site, sous réserve de validation par l'assemblée générale annuelle

- Bureaux du laboratoire

Afin d'aboutir à un fonctionnement efficace, il a été décidé d'instituer un bureau par site :

- un directeur (Orléans) et un directeur adjoint (Tours) – cette répartition valant pour ce contrat, sans engagement pour les suivants.
- un responsable des finances et des moyens par site ayant pour mission de suivre les entrées et les dépenses sur chaque site
- un représentant des filières de formation (parcours recherche des Masters et doctorats) par site
- un chargé des activités scientifiques (séminaire, colloques...) par site,
- un représentant des doctorants par site,
(en cas d'affectation de personnel., un représentant des IATOS sera ajouté au Conseil)

- Directeur

Le Directeur, et le Directeur-adjoint qui lui correspond sur l'autre site, ont la responsabilité de l'animation du laboratoire, de son activité et du respect de ses règles. Dans chacun des sites, ils représentent le laboratoire auprès des instances universitaires et des partenaires. Seul le Directeur est habilité à représenter l'ensemble du laboratoire. Ils sont désignés par élection de l'assemblée générale pour la durée du contrat.

- Responsables d'équipe

Les responsables d'équipes sont élus par les membres permanents du laboratoire engagés dans l'équipe concernée

7. Budget

- Sources : outre le budget récurrent alloué par le Conseil Scientifique, toute source contractuelle accessible
- Répartition : selon les besoins et les disponibilités, dans le respect des engagements contractuels

Gestion : les fonds sont alloués par le bureau de site après demande motivée, puis laissés à la libre utilisation de l'équipe ou du chercheur concerné sous réserve de justification. La répartition du budget est soumise à l'assemblée.

LISTE DES LABORATOIRES CNRS

Section 34 / INSHS

Recension établie au 1^{er} janvier 2008
 SHS département principal / section 34 principal ou secondaire
 En gras, les UMR centrées sur les sciences du langage ; rayées, celles qui ont été dés-UMRisées.

ILE-DE-FRANCE

26 laboratoires dont 13 UMR centrées sur les SDL (1 supprimée) = 12 UMR

Laboratoire de Phonétique et Phonologie	Jacqueline VAISSIERE	LPP	PARIS
Structure formelle du langage : typologie et acquisition, métrique et poétique	Clive PERDUE	SFL	ST DENIS
Langues et civilisations à tradition orale	Zlatka GUENTCHEVA- DESCLES	LACITO	VILLEJUIF
Laboratoire de linguistique formelle	Alain KIHM	LLF	PARIS
Modèles, Dynamiques, Corpus	Bernard LAKS	MODYCO	NANTERRE
Laboratoire d'ethnologie et de sociologie comparative	Anne-Marie PEATRIK		NANTERRE
Lexiques, dictionnaires, informatique	Salah MEJRI	LDI	VILLETANEUSE
Histoire des théories linguistiques	Sylvie ARCHAIMBAULT	HTL	PARIS
Centre de recherche en épistémologie appliquée	Paul BOURGINE	CREA	PARIS
Langues, textes, traitements informatiques, cognition	Laurence DANLOS	LATTICE	MONTROUGE
Langues – Musiques Sociétés	Frank ALVAREZ-PEREYRE	LMS	VILLEJUIF
Institut Jean-Nicod	Pierre JACOB	IJN	PARIS
Institut des Textes et Manuscrits Modernes	Pierre-Marc DE BIASI	ITEM	PARIS

Centre d'études des langues indigènes d'Amérique	Francisco QUEIXALOS	CELIA	VILLEJUIF
Langage, langues et cultures d'Afrique noire	Martine VANHOVE	LLACAN	VILLEJUIF
Centre de recherches linguistiques sur l'Asie orientale	Redouane DJAMOURI	CRLAO	PARIS
Institut d'Histoire et de Philosophie des Sciences et des Techniques	Jacques DUBUCS	SHS	PARIS
Cultures, Langues, Textes	Isabelle PASTOR-SOROKINE		VILLEJUIF
TGE-ADONIS	Gilbert PUECH		PARIS
Relais d'information sur les sciences de la cognition (RISC)	Jean LORENCEAU		PARIS
Institut de linguistique française (ILF)	Christiane MARCHELLO-NIZIA	FR	PARIS
Typologie et universaux linguistiques : données et modèles	Stéphane ROBERT	FR	PARIS
Communication et politique	Isabelle VEYRAT-MASSON		PARIS
Description et modélisation en morphologie	Bernard FRADIN	GDR	PARIS
Sémantique et modélisation (SEM)	Francis CORBLIN	GDR	PARIS
Linguistique latine	Michèle FRUYT	GDR	PARIS

PROVINCE ET OUTRE-MER

16 laboratoires dont 12 UMR centrées sur les SDL (3 supprimées) = 9 UMR

Interactions, Corpus, Apprentissage, Représentations	Lorenza MONDADA	ICAR	LYON
Cognition, Langues, Langages, Ergonomie	Jacques DURAND	CLLE-ERS	TOULOUSE
Laboratoire PRAXILING	Jacques BRES	PRAXILING	MONTPELLIER
Centre de recherches sur la langue et les textes basques	Bernard OYHARCABAL	IKER	BAYONNE
Dynamique du langage	François PELLEGRINO	DDL	LYON
Bases corpus et langage	Tobias SCHEER	BCL	NICE
Laboratoire Parole et Langage	Philippe BLACHE	LPL	AIX EN PROVENCE
Centre d'Etudes Supérieures de la Renaissance	Marie-Luce DEMONET	CESR	TOURS
Moyen-Âge	Patrick CORBET	SHS	NANCY
Analyse et traitement informatique de la langue française	Jean-Marie PIERREL	ATILF	NANCY
Langues, textes et communications dans les espaces créolophones et francophones	Lambert-Félix PRUDENT	SHS	ST DENIS
Savoirs, Textes, Langage	Fabienne BLAISE	STL	VILLENEUVE D ASCQ
UMS de la Maison de la recherche de l'université de Toulouse le Mirail	Bertrand JOUVE	UMS	TOULOUSE
Dynamiques sociolinguistiques	Régine DELAMOTTE- LEGRAND	DYALANG	MONT ST AIGNAN
Centre de Recherches inter- langues sur la signification en Contexte	Jacques FRANCOIS	CRISCO	CAEN
Approche pluridisciplinaire de la production verbale écrite	Denis ALAMARGOT	GDR	POITIERS

LABORATOIRES EN RATTACHEMENT SHS SECONDAIRE

ILE-DE-FRANCE

6 laboratoires

SHS département secondaire / section 34 secondaire

Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI)	Patrick LE QUERE	ST2I	ORSAY
Institut Jean le rond d'Alembert	Gérard MAUGIN	ST2I	PARIS
Vulnérabilité, adaptation et psychopathologie	Roland JOUVENT	SDV	PARIS
Centre d'analyse et de mathématique sociale (CAMS)	Henri BERESTYCKI	ST2I	PARIS
Laboratoire des Usages en Technologies d'Information Numériques (LUTIN)	Dominique BOULLIER	ST2I	PARIS
Laboratoire psychologie de la perception	John Kevin O'REGAN	SDV	PARIS

PROVINCE ET OUTRE-MER

9 laboratoires

SHS département secondaire / section 34 secondaire

GIPSA	Jean-Marc CHASSERY	ST2I	ST MARTIN D'HERES
Laboratoire d'Informatique de Grenoble (LIG)	Brigitte PLATEAU	ST2I	ST MARTIN D'HERES
Institut de recherche en Informatique de Toulouse (IRIT)	Luis FARINAS DEL CERRO	ST2I	TOULOUSE
Laboratoire de psychologie sociale et cognitive (LAPSCO)	Michel FAYOL	SDV	CLERMONT FERRAND
Groupe de Recherche en Informatique, Image, Automatique et Instrumentation de Caen (GREYC)	Régis CARIN	ST2I	CAEN
Laboratoire d'informatique Fondamentale de Marseille (LIF)	Paul SABATIER	ST2I	MARSEILLE
Groupe d'imagerie neuro-fonctionnelle	Bernard MAZOYER	SDV	CAEN
Laboratoire Langage, Mémoire et Développement Cognitif (LMDC)	Jean-Francois ROUET	SDV	POITIERS
Laboratoire lorrain de recherche en informatique et ses applications (LORIA)	Karl TOMBRE	ST2I	VANDOEUVRE LES NANCY

LA RECHERCHE A LA BNF

(version du 25 juin 2010 - extraits)

site Internet : <http://www.bnf.fr/fr/>

La recherche subventionnée

Principalement cofinancée par la BnF et par le département de l'Enseignement supérieur, de la recherche et de la technologie du ministère de la Culture et de la Communication, la recherche subventionnée constitue une part très importante de l'activité de recherche au sein de la Bibliothèque. Elle répond à des exigences scientifiques précises en termes d'obligation de résultats, de travail en partenariat et de programmation rigoureuse dans le temps.

Elle est organisée dans le cadre de plans triennaux, dont celui en cours porte sur les années 2010-2012. Cette organisation remonte à 1995 et le nouveau plan lancé est donc le sixième de ce type. Au cours de cette période, l'accent est mis sur l'évaluation des programmes au sein de comités scientifiques particuliers et du Comité de la Recherche de la BnF.

La recherche non subventionnée

À côté des programmes de recherche subventionnés, l'activité de recherche non subventionnée est également essentielle et responsable en grande partie du rayonnement des différents départements de la Bibliothèque. Elle relève davantage de l'activité scientifique courante de ces derniers et est financée sur leur budget.

Informations sur les départements

Les travaux scientifiques menés répondent à l'une des missions essentielles de la Bibliothèque nationale de France :

- * leurs champs d'investigation sont, par essence, aussi vastes et divers que les collections de la bibliothèque elle-même

- * leurs niveaux d'ambition sont également divers (du dénombrement ou signalement d'objets au catalogue raisonné...) et exigent des travaux de quelques années à plusieurs générations...

- * leur publication s'effectue sous des modalités très différentes (catalogues de collections ou d'expositions, répertoires, bibliographies, publications savantes, articles de revue, communications dans des manifestations universitaires souvent internationales...)

- * enfin, dans le domaine de compétence et d'expertise qui est le leur, certains conservateurs de l'établissement ont acquis aujourd'hui une envergure comparable à celle de prédécesseurs comme Léopold Delisle, Ernest Babelon ou Jean Adhémar, qui ont marqué l'histoire de la recherche et de l'érudition. Ces travaux contribuent de façon essentielle au rayonnement de la bibliothèque.

Axes de recherche

Qu'ils soient subventionnés ou non, les programmes de recherche se développent selon les mêmes axes :

- * Bibliographie : production d'outils pour la recherche : inventaires, catalogues raisonnés, répertoires... ;

- * Conservation : recherche de nouveaux procédés de conservation aussi bien des supports « papier », numériques, microformes... que de leur contenant, à savoir boîtes d'archivage, pochettes...

- * Histoire du livre : étude et production d'outils concernant l'histoire du livre ;

- * Iconographie : étude méthodologique des représentations plastiques que la Bibliothèque conserve (peintures y compris enluminures, gravures, photographies quel que soit leur support, affiches, dessins, etc.) avec leurs sources, significations et classements ;

- * Patrimoine de la musique et du spectacle : création de catalogues thématiques et de répertoires, recensement de sources musicales et d'arts du spectacle ;

- * Numismatique : études sur les collections de monnaies du département des Monnaies, médailles et antiques en relation avec les autres cabinets numismatiques français et étrangers ;

- * Supports numériques et nouvelles technologies : mise au point d'outils de structuration des documents électroniques ;

- * Supports, usages et circulation de l'écrit : étude des supports d'information autres que le livre imprimé sur le plan matériel et intellectuel.

Evolution

La recherche s'inscrit aujourd'hui dans ces perspectives d'évolution :

- * un effort de resserrement des programmes,

- * un rééquilibrage des travaux de recherche entre les départements s'appuyant sur une longue tradition et des départements dont les domaines sont plus récents,

- * la mise en place d'une programmation, l'amélioration du suivi, le développement du faire savoir et de la valorisation,

- * le développement de partenariats et une plus grande ouverture vers le monde de la recherche, en particulier par le biais de conventions avec le Centre National de la Recherche Scientifique et les unités mixtes de recherche (UMR), de l'appel à chercheurs lancé chaque année auprès des jeunes doctorants français et étrangers.

Les programmes de recherche font l'objet de publications scientifiques.

Bibliothèque nationale de France
Plan triennal de la recherche 2010 – 2012

Bibliographie

In Principio 2	2012	Manuscrits
Catalogue des manuscrits enluminés d'origine germanique	2012	Manuscrits
Inventaire des incunables scientifiques de la bibliothèque de l'Arsenal	2012	Bibliothèque de l'Arsenal
Un cabinet savant à l'époque des Lumières : Jean-Baptiste d'Anville	2013	Cartes et plans
Les portulans conservés en France	2012	Cartes et plans

Conservation

Conservation des documents audiovisuels 2012		Audiovisuel
--	--	-------------

Histoire du livre

Renouard. Imprimerie et librairie parisienne du XVI ^e siècle	s.d.	Réserve des livres rares
--	------	--------------------------

Iconographie

Mise en valeur des fonds d'architectes	2012	Estampes et photographie
Fonds des architectes aux XIX-XX ^e s.	2012	Mission pour la gestion de la production documentaire
Manuscrits japonais à l'époque d'Edo	2012	Manuscrits

Numismatique

Catalogue des monnaies de l'Empire romain	2012	Monnaies, médailles et antiques
Trésors Monétaires	s.d.	Monnaies, médailles et antiques
Catalogue des monnaies celtiques	2012	Monnaies, médailles et antiques

Patrimoine de la musique et du spectacle

Catalogue thématique de l'œuvre
de Jean-Philippe Rameau s.d. Musique

Répertoire international des sources
musicales (R.I.S.M.) s.d. Musique

Supports numériques et nouvelles technologies

Caractérisation des Documents
Numérisés et prédiction de performances
des techniques d'OCR et d'Indexation 2013 Conservation

Supports, usages et circulation de l'écrit

Reliures françaises des collections
de la BnF, du Moyen Âge au XIX^e siècle 2012 Réserve des livres rares

Convention CNRS – BnF

La convention signée le 2 mars 2005 par le président de la BnF, Jean-Noël Jeanneney, et Bernard Larrouturou, directeur général du CNRS, reconduit pour une nouvelle période de cinq ans la coopération engagée en 1997 par les deux institutions autour de la production, de l'usage et de la conservation des supports d'information.

L'objectif est de poursuivre ensemble des programmes de recherche variés : fabrication, histoire et conservation du papier, histoire du livre, numismatique, patrimoine musical français, bibliographies spécialisées, traitement automatisé du langage, production d'outils d'accès aux sources documentaires, nouvelles technologies de l'information et de la communication (NTIC)...

La coopération entre les deux institutions consiste en un échange de compétences et de savoir-faire de leur équipes respectives. Elle passe aussi par la mise à disposition de locaux et de matériels ainsi que par la gestion et la valorisation concertée des résultats issus de ces recherches.

Cette nouvelle convention-cadre marque la volonté commune du CNRS et de la BnF de renforcer encore leur partenariat, engagé depuis huit ans.

Les programmes communs de recherche CNRS-BnF
Sept unités mixtes de recherche du CNRS sont particulièrement concernées par cette coopération scientifique avec la BnF :

- le Centre de recherches sur la conservation des documents graphiques (CRCDG – Ministère de la culture / MNHN / CNRS),
- l'Institut de recherche et d'histoire des textes (IRHT),
- l'Institut d'histoire moderne et contemporaine (IHMC – CNRS / ENS)
- le Centre d'Études Préhistoire, Antiquité, Moyen-Age (CEPAM – CNRS / Université de Nice),
- l'Institut de recherche sur le patrimoine musical en France (IRPMF – CNRS / MCC / BnF)
- le Centre de Recherche et de Restauration des Musées de France (C2RMF – CNRS / MCC)
- l'Institut de Recherche sur les Archéomatériaux (IRAMAT – Université Bordeaux 3 / Université d'Orléans / CNRS)).

Le CRCDG et la BnF coopèrent sur deux programmes scientifiques : l'étude de l'influence des traitements de restauration sur la conservation des documents papiers comportant des encres ferrogalliques et la recherche d'un cuir de tannage végétal pour la conservation des livres et reliures anciennes dégradées.

L'IRHT et la BnF vont poursuivre leur collaboration autour de la description, l'étude et le catalogage des manuscrits médiévaux et des premiers imprimés qui prendra fin en 2007. Le partenariat commun développe des projets de co-édition, la formation des personnels ainsi que la diffusion et la valorisation du patrimoine culturel et de la recherche scientifique (production et diffusion de revues communes par des liens ou des portails informatiques...).

L'IHMC et la BnF continuent leur partenariat pour la réalisation en commun de la *Bibliographie annuelle de l'histoire de France*.

Le CEPAM et la BnF coopèrent jusqu'en 2006 sur le Catalogue des monnaies impériales. L'objectif est d'étudier, de recenser et de reclasser les monnaies impériales des collections du Cabinet des médailles de la BnF ainsi que de constituer une base de données et de publier un catalogue.

L'IRPMF et la BnF poursuivent leur collaboration sur le programme RISM « Répertoire et inventaire des sources musicales » qui consiste à inventorier et décrire des manuscrits musicaux appartenant à la période 1570-1800, conservés dans les bibliothèques françaises. Il s'agit de continuer le travail engagé de catalogage des manuscrits musicaux et de procéder à sa publication.

Le C2RMF mène des recherches conjointes avec le CRCDG et les ateliers de restauration de la BnF.

Enfin, le Centre Ernest-Babelon (de l'IRAMAT) et la BnF collaborent depuis de nombreuses années dans deux domaines, la numismatique et la codicologie. En numismatique, plusieurs axes de recherches sont conduits par les deux organismes : la réalisation de catalogues de monnaies gauloises et la caractérisation des métaux monnayés, en vue d'études dans le domaine de l'histoire économique et monétaire. En codicologie, le partenariat porte sur l'étude des manuscrits latins et arabes, en se fondant sur la caractérisation des encres et des pigments et colorants des enluminures.

Pour la constitution d'une UMR BnF / LLL / CNRS
- CODEX -
(Centre d'Opérabilité des Données pour l'Exploitation des Corpus Sonores)

Un objet scientifique : le corpus sonore

Si les corpus sonores ont émergé il y a une centaine d'années comme de possibles moyens de recherche à la disposition des sciences humaines et sociales, il a fallu attendre la diffusion du magnétophone pour qu'il soit possible de recueillir aisément de la parole sur le terrain. Il en est résulté un accroissement des enregistrements disponibles tel que seule l'informatique peut en rationaliser l'exploitation. Ainsi, en une vingtaine d'années, la référence à des sources contrôlables est devenue inhérente à l'argumentaire scientifique des linguistes, des anthropologues, des sociologues... C'est à cette pratique, sinon nouvelle du moins récente, qui conditionne largement le travail de laboratoire, que répond le champ d'investigation circonscrit par cette UMR avec une double spécificité :

- d'*objet*, puisqu'elle centre les investigations sur une classe particulière de corpus : ceux qui réunissent des matériaux langagiers pérennisés, conservés avant tout dans leur forme sonore, après qu'ils ont été établis ou collectés à des fins d'étude et de transmission, avec l'objectif de rendre compte de leurs propriétés sans obérer leur conservation ;
- de *méthode*, en unifiant des appropriations fragmentaires par une maîtrise de l'ensemble du processus qui constitue un corpus, de sa conception à sa consultation, avec une attention particulière concernant les procédures d'archivage.

Bien que l'étude soit volontairement restreinte aux documents sonores, trois types d'éléments seront associés dans l'inventaire :

- les enregistrements sous la forme où ils ont été réalisés et/ou conservés, ainsi que les duplications effectuées successivement,
- les annotations, qu'elles soient intégrées dans l'enregistrement (sous forme de plages ou de commentaires insérés) ou qu'elles soient adjointes sur des supports graphiques au moyen de transcriptions (orthographiques, phonétiques...), d'indications situationnelles et interactionnelles, etc.
- les outils et les systèmes de description qui assurent la contextualisation par l'assignation des locuteurs ou de la situation, par l'indexation des contenus, mais aussi qui préservent la maîtrise des techniques de saisie, de gravure et de lecture de l'enregistrement, un corpus sonore étant indissociable des dispositifs qui en assurent la restitution.

Les corpus ainsi entendus correspondent à des produits conçus à l'intérieur d'une pratique scientifique pour être éprouvés par des analyses et pour se prêter à des applications. Ils peuvent devoir leur existence à un programme qui les a constitués ou bien résulter d'une sélection dans des fonds préexistants. Quelles que soient les sollicitations, les communautés participant à l'exploitation du document sonore ont pour premier souci la préservation du matériel recueilli, ce qui signifie non seulement une exploitation non destructive des ressources mais aussi l'actualisation continue de l'accès aux contenus, pour surmonter les conséquences de transformations technologiques qui renouvellent constamment les formats et les appareils de lecture, et pour compenser la dégradation du contenant et du contenu des données. La pérennisation des documents suppose la définition d'une méthodologie dont les principes ont été définis par les acteurs eux-mêmes dans *Corpus oraux, Guide des bonnes pratiques*, sous la direction d'O. Baude (2006).

Deux approches complémentaires

L'accroissement des réserves de documents sonores et leur prise en charge patrimoniale ont mis en évidence la fragilité des supports et l'obsolescence des systèmes de restitution. Les fonctions de collecte et de conservation, de stockage et d'analyse, de balisage et d'indexation, de mise à disposition et d'application, effectuées séparément dans les laboratoires et les centres d'archivage, ne correspondent plus aux exigences induites par l'évolution des technologies et des pratiques scientifiques. La création d'une UMR, une *unité mixte de recherche* au sens plein du terme, prend acte qu'une division du travail, qui a prévalu jusqu'à maintenant, ne se justifie plus et qu'il faut établir de nouvelles formes d'expertise et de connaissance par l'addition des expériences et des compétences.

Alors que le document sonore a été appréhendé jusqu'à présent comme un objet, transférable de l'équipe qui le recueille à l'institution qui le conserve, le programme de l'UMR conçoit les corpus dans leur dynamique, c'est-à-dire comme un processus, un enchaînement d'opérations solidaires, depuis les hypothèses théoriques sous-jacentes à une décision de collecte jusqu'à la communication au public. Dans l'intervalle, le document sonore demande à être constamment retravaillé afin que l'archive demeure comme un témoignage vivant.

D'abord, son contenu peut être rattaché à des séries indéfiniment extensibles. Si même il se trouve catalogué dans une collection pour laquelle l'enrichissement est tenu pour achevé, ses modalités de conservation (codage et formatage) continuent d'accompagner sa transmission. Surtout, les annotations qui en assurent le traitement ne peuvent s'interrompre dans une évolution qui, en contribuant à l'enrichissement continu du corpus, en modifie l'accès et les lectures.

Dans cette perspective, la collaboration entre les parties contractantes se pense dans une solidarité des gestes scientifiques, une complémentarité des objectifs et des modes de fonctionnement, avec des recoupements et des relais qui confortent la pertinence du projet, sa cohérence et plus encore sa nécessité. C'est ce qui ressort du circuit du document dont on reconstituera le parcours comme indiqué dans le tableau :

Légendes et sigles

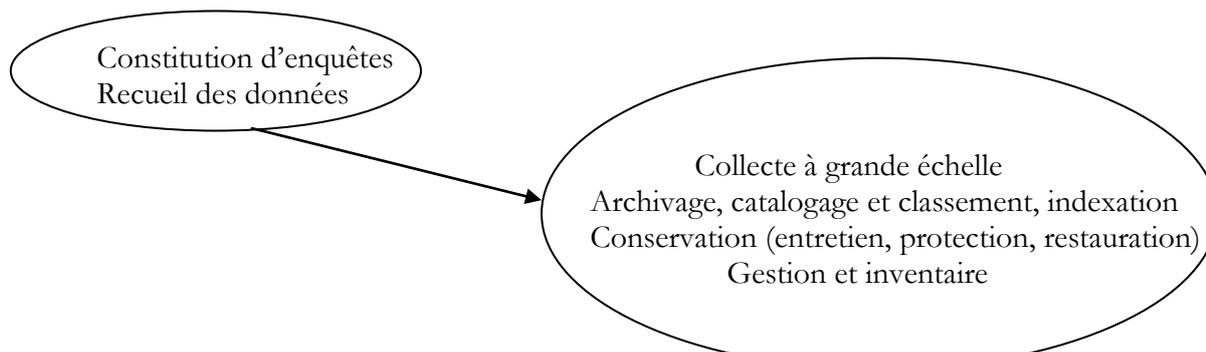
- UMR : Unité Mixte de Recherche (Université / Grand Etablissement / CNRS)
- EA : Unité Propre de Recherche à l'Enseignement Supérieur / Equipe d'Accueil
- CORAL : Centre Orléanais de Recherche en Anthropologie et Linguistique
- LLL : Laboratoire Ligérien de Linguistique
- BnF : Bibliothèque Nationale de France (Ministère de la Culture)

*Circuit du document et phases de traitement
selon les compétences des parties contractantes dans l'UMR*

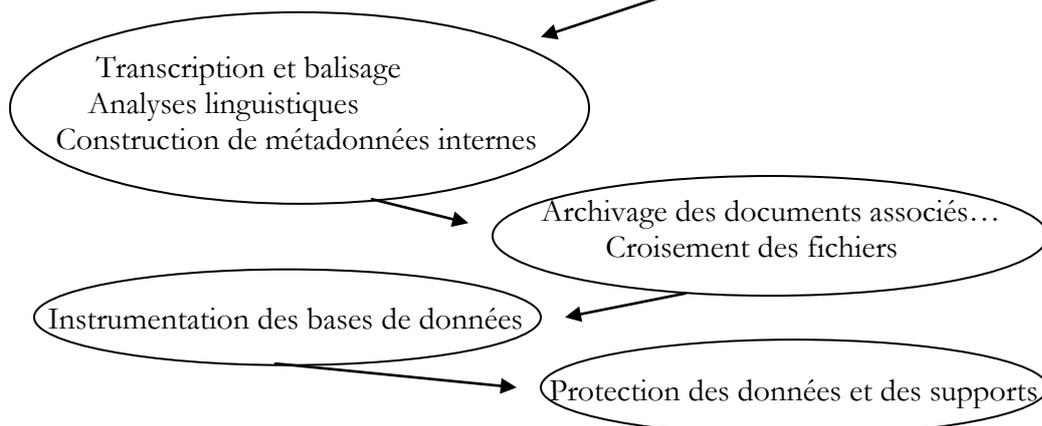
EA LLL / U. d'Orléans et Tours

BnF / Département de l'Audiovisuel
Service des Documents Sonores

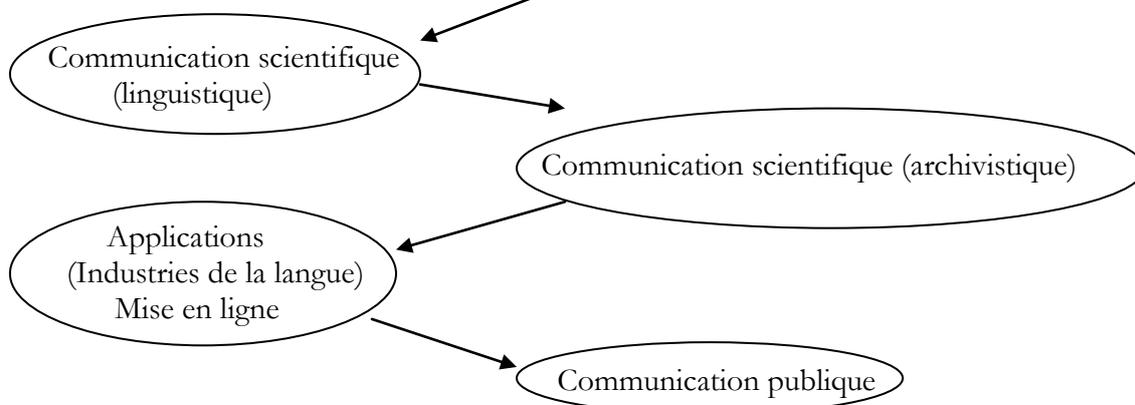
1. Création de corpus



2. Exploitation des corpus



3. Diffusion des corpus



A partir de cette circulation entre deux institutions, qui représentent aussi deux types de compétences qu'il s'agit de réunir, l'une des perspectives de l'UMR est de parvenir à produire des usages différenciés et multiples des corpus. En lieu et place d'un document collecté par un laboratoire pour une seule exploitation finalisée, l'objectif serait de rendre possible une réutilisation des données. Cela suppose d'engager un traitement d'ensemble, qui suppose des modèles et des références, mais aussi un accroissement des capacités d'interopérabilité à tous les niveaux d'accès et d'analyse.

Dans cette conception, le travail d'un laboratoire comme le CORAL ou LLL ne représente qu'une fraction de ce que le Service des Documents Sonores du Département de l'Audio-Visuel de la BnF doit traiter mais il peut avoir valeur de test pour la définition concertée d'une politique raisonnée de l'archive sonore. Qu'il s'agisse des travaux effectués à Orléans, ESLO 2 ou l'Encyclopédie palikur par exemple, ou de collectes effectuées par d'autres laboratoires, l'accompagnement à ses différentes phases de l'élaboration d'un corpus permet aux conservateurs et aux personnels de documentation de suivre *in situ* les segments de la fabrique des données sonores auxquels ils n'auraient pas accès et aux chercheurs de profiter de leurs compétences en matière de saisie, d'indexation ou d'exploitation. Dans tous les cas, les études ne sauraient avoir vocation à se limiter aux apports de l'équipe.

Dans cette collaboration, la contribution attendue du CNRS s'apparente à celle dont bénéficient les équipes qui se sont mises en place autour des documents écrits, que ce soit dans le cadre d'un partenariat avec la BnF, :

- l'Institut de Recherche et d'Histoire des Textes (IRHT) et
 - le Centre de Recherches sur la Conservation des Documents Graphiques,
- ou dans des conventions avec les Universités :
- Analyse et Traitement Informatique de la Langue Française (ATILF, UMR CNRS/Nancy 2)
 - Institut des Textes Et Manuscrits modernes (ITEM, UMR CNRS/ENS)

Le spectre d'activités couvert par ces laboratoires, en matière d'archives, n'inclut pas les ressources orales. Il s'agit donc d'assurer une extension des champs documentaires, confinés jusqu'à présent dans les *scripta* et les *realia*, afin d'y inclure les *vocalia*, de les rendre accessibles à la recherche en rapprochant le traitement des corpus sonores des modalités sanctionnées par le travail des équipes qui se consacrent aux recueils de textes et de manuscrits. C'est l'occasion d'opérer un rapprochement entre l'enquête et l'archivistique à l'intérieur d'un domaine qu'elles partagent sans se rencontrer, en dépit d'initiatives associatives comme l'AFAS (Association Française des Détenteurs de Documents Audiovisuels ou Sonores) ou de projets comme le CRDO (Centre de Ressources pour la Description de l'Oral).

Linguistique et traitement du document sonore

La synergie attendue de l'organisation en UMR répond à quatre objectifs :

- expliciter les critères de constitution de corpus, de traitement, d'archivage...
- représenter sur un mode déclaratif les connaissances qui sous-tendent les catégorisations effectuées par les balises ou les procédures d'indexation,
- modéliser les procédures d'archivage, avec intégration des liens entre les différents documents constitutifs du corpus, notamment les méta-données,
- contribuer à élaborer une notation formelle afin d'exercer un retour réflexif sur les documents pour en contrôler le contenu et faciliter l'accès aux contextes situés.

Le traitement du document sonore est conditionné par une conception du corpus qui a pour ambition de ne pas restreindre son usage à ceux qui l'ont produit mais qui anticipe l'accessibilité auprès d'autres communautés, celles des chercheurs français et étrangers, mais aussi, pour certains enregistrements, à destination du grand public. Cela suppose que l'archive soit paramétrée de façon à assurer la consultation par différents systèmes informatiques, sur des supports fixes ou à distance, et que la structure d'interrogation soit conçue en fonction des normes, standards et usages pratiqués selon le type de diffusion choisi et les destinataires potentiellement concernés. L'impératif d'interopérabilité implique une confrontation des requêtes formulées à l'ensemble des sources en catalogue et à la diversité de leurs formats.

On postule que l'archive sonore constitue le document pivot autour de quoi s'organisent les documents associés et l'ensemble du travail du conservateur et du chercheur. Pour être exploitable et disponible, à tous les moments de sa consultation, en particulier dans une perspective patrimoniale de transmission, il est prévu la conservation des protocoles et des outils qui en assurent le stockage, le transport et la lecture dans un format accessible de façon autonome. La croissance des données et l'évolution rapide des techniques qui assurent le conditionnement des documents sonores, aussi l'imprévisibilité des demandes futures de consultation, tant dans leur approche que dans leur mode opératoire, induisent la poursuite d'une réflexion pour que les compétences de manipulation et d'interrogation de corpus sonores soient portées au niveau atteint par les études sur les corpus textuels.

Ainsi, le point d'application d'une recherche engagée en commun ne concerne ni la strate informatique au principe des langages électroniques et de la documentation automatique, ni la gamme des exploitations assurées par les destinataires ultimes qui peuvent être issus de différentes disciplines : historiens, sociologues, anthropologues, linguistes... Le but n'est pas de se substituer aux communautés scientifiques mais de les accompagner dans l'ouverture de leurs données à d'autres sollicitations, de dépasser la correspondance terme à terme d'un corpus sonore et d'une enquête thématique.

Les études programmées dans le cadre de l'UMR se concentreront donc sur l'interface entre les matériaux traités comme archives – supports, discours enregistrés et méta-données –, et une élaboration qui met à contribution les sciences du langage pour augmenter l'accessibilité et l'enrichissement des corpus traités. Elles se prolongeront jusqu'à l'observation des usages et des exploitations dont la connaissance est indispensable à la modélisation des procédures.

L'accès aux contenus

On part du principe que la reconnaissance du corpus est réalisée au moyen d'un système de descripteurs et de relations, représentés par des mots ou des algorithmes et organisés de façon hiérarchique. Leur efficacité est appréciée en fonction du degré de réduction du temps d'accès à l'information définie comme pertinente par la requête. Par ailleurs, une entité n'est constituée comme telle dans le corpus qu'à partir du moment où elle a été l'objet d'une description et d'une indexation. La structuration ne porte pas directement sur les données mais sur les informations qui leur sont associées, descripteurs formalisés et entités taxinomiques organisées en « métadonnées ».

C'est sur des systèmes de représentation et d'indexation que peuvent être établis une terminographie, des index, des thesaurus, et même aujourd'hui des « ontologies ». Quelle que soit l'approche privilégiée, elle demeure tributaire d'une connaissance approfondie des corpus pour laquelle les procédés d'enregistrement et de conservation et la définition des objets éligibles sont

des pré-requis, sachant que les représentations sont construites en fonction des outils de traitement qui déterminent les modalités de l'indexation.

La finalité de la collaboration réside dans cette conjonction d'une signalétique, réalisée par l'archiviste, et d'une analyse du document effectuée à partir de ses caractéristiques langagières. Le langage XML, qui fonctionne aujourd'hui comme le standard retenu par la plupart de ceux qui s'engagent dans le balisage des corpus sonores, assure l'interface entre requêtes et contenu des données. De la convergence entre l'exploration du contenu et la définition signalétique procède la modélisation de méthodes efficaces aussi bien pour la création que pour l'exploitation des corpus. En particulier, à partir d'un descripteur formulé au moyen d'un ou plusieurs termes extraits des données documentaires et/ou des transcriptions, on peut définir les « pointeurs » qui retrouvent les séquences du fichier sonore ou les séquences de texte correspondant à la recherche formulée.

La contribution de l'UMR au développement international des ressources électroniques sur les langues la désigne comme un partenaire important de l'*Open Language Archives Community* (OLAC) pour la création, en français et dans d'autres langues, d'une phonothèque virtuelle mondiale, en cours de développement collectif dans un cadre associatif. Après la réaction concertée des bibliothèques européennes pour la protection et la mise à disposition de leur patrimoine écrit, c'est à une initiative d'envergure comparable qu'il est prévu d'apporter une dimension de recherche et une visibilité sur la Toile.

Maniabilité des corpus sonores : contraintes et solutions

L'exploitation rationnelle des masses de données disponibles suppose d'automatiser et de fiabiliser la structure et les liens inscrits du corpus. Au centre des relations ordonnées entre les différents supports, il y a les corrélations établies entre les fichiers sonores et les fichiers de transcription correspondants, dont le format informatique diffère, rendant plus complexe l'interconnexion entre les pages du fichier sonore et la transcription textuelle qui la convertit en chaînes de caractères. La pluralité de formats de stockage (WAV / MP3, ASCII / XML) accroît les risques d'impédance quand les enregistrements n'ont pas été effectués en suivant les prescriptions.

Un étalonnage chronométrique permet d'assurer un premier repérage, à la fois homogène et indistinct. Il est intégrable à une copie de l'enregistrement, soit par des découpes en pages, soit par des signaux intercalés (pas forcément perceptibles à l'écoute). Dès lors que la valeur temporelle reste déterminante, en tant qu'elle est la seule à être universelle et fixe, il importe que le calage de la durée, l'invariabilité du défilement et la stabilisation du format soient assurés.

Un fichier sonore ne vaut, pour le chercheur, que par sa disponibilité effective et par la possibilité de l'interroger à partir de ses transcriptions et des métadonnées. L'accès au document procède des documents reliés : connaissance de la littérature grise, formalisation des métadonnées exploitation de la source par des travaux de laboratoire et des publications, systèmes de renvois, descripteurs. La massification des données et l'automatisation des procédures informatiques requièrent une sophistication des relations entre le document pivot et ses documents annexes dont la relation est rendue d'autant plus complexe que sera prise en compte la mémoire des états du corpus.

L'ensemble des acteurs contribuant à la fabrication du corpus (de la conception à la diffusion) s'accordent sur le rôle que jouent l'archivage et l'exploitation du document sur sa conception et sa

représentation. Après qu'ont été exposées les attentes sociolinguistiques, la définition des terrain, la sélection des témoins et la constitution des enquêtes, les métadonnées conditionnent la teneur et la structuration des corpus, leur degré de transparence ou au contraire d'opacité aux sollicitations dont ils sont l'objet. Un déplacement intervient qui conduit à distinguer, au centre des requêtes, l'archive sonore comme document référentiel d'un côté et, de l'autre, les annotations textuelles qui, au cours de leur évolution, sont établies comme document d'accès. Or, les outils d'indexation, de classement, de recherche et de gestion, quels que soient leur principe et leur finalité, sont toujours en dernière instance des modes linguistiques d'organisation de la connaissance, en particulier des éléments du lexique des langues naturelles, dans leur distribution, leur hiérarchie et leur agencement interne. A ce titre, les sciences du langage sont, par définition, contributives pour la conservation et l'exploitation des corpus sonores.

Objectifs et moyens de l'UMR

Grâce au rassemblement de nos compétences, nous pourrions déterminer les apports d'une démarche linguistique dans la gestion des corpus avec la conviction que poser des questions linguistiques permet d'anticiper les procédures de conservation. L'objectif est de modéliser les démarches : un savoir acquis lors la création d'un corpus doit pouvoir être réexploité de façon contrôlée sur des archives du même type et contribuer à l'élaboration de protocoles renouvelables.

L'enjeu est double. D'abord, enrichir les corpus. Ensuite, prolonger leur pérennité en conservant les savoirs et les expertises à même de les valoriser et de les lire. Ainsi, la généralisation de l'usage du magnétophone dans les pratiques scientifiques, il y a une quarantaine d'années, a coïncidé avec un cycle d'activité de chercheurs et de techniciens. Il nous appartient de veiller à la conservation et à la disponibilité des données qu'ils ont constituées, comme le LLL en a fait la démonstration à partir du corpus ESLO.

À partir de la numérisation, on réaccorde les deux approches complémentaires qui ont fait des corpus un des outils fondamentaux des sciences humaines et sociales : l'archivistique et l'éditionnalisation du contenu par son traitement linguistique, en mutualisant l'expertise des analyses sur l'expression et le contenu et le traitement documentaire des données.

Document de travail (2007-2009)

CAT COD

Appel à colloque - 2008

The number of oral recordings which are digitized and available for the study and the description of languages remained quite small for a long time and they were confined within specialized agencies which distribute them. With the development of the Web and of the storage, diffusion and processing capacities, it is now affordable for smaller structure (e.g. research laboratories) to distribute themselves their oral resources.

We are henceforth at a key stage where the capitalization of the efforts of linguistic data digitalization becomes possible, this in order to form the empirical basis for new research projects.

The observation of the coding and cataloguing standards of this type of resources within the various existing projects in France and Europe shows a great heterogeneity among practices. Whereas these very projects had been set up with the aim of facilitating the access, the sharing or the preservation of the data, we notice that the diversity within the formats, the codings and the protocols used, precisely limits these objectives.

In this symposium, we wish to report on the major initiatives within the field of digital documents management which would potentially have an important influence on standardization. To do so, we would like to stress two specific aspects which are: the cataloguing of an oral resource and its coding.

Cataloguing

The recent and fast growth of the number of oral recordings available on the Web needs to be accompanied by an important effort of description and referencing in order that these data should be easily accessible, should not be buried in the mass of available data and that their management should be facilitated.

Thus, it is necessary to distinguish the cataloguing practices whose purpose is to follow up the life cycle of a project of constitution of resources from those whose purpose is to guide the exploitation or the preservation and the distribution of these resources.

This cataloguing activity is all the more important and urgent as the handling operations of this mass of data increase as well (exchanges, maintenance, research, etc.). Some communities already got themselves strongly organised and have sometimes established standards such as the Dublin-Core as far as the resources on the Web are concerned, or dating from an earlier period, the MARC standards maintained by the Library of Congress in order to describe bibliographical resources. Finally more recently, smaller linguistic communities have established cataloguing proposals (OLAC, IMDI). People have now had enough experience in the use of these proposals to be able to criticize them, propose some improvements, give rise to new thinking and establish a minimal charter that should be respected by those who wish to get involved in the distribution of linguistic oral resources in order to facilitate their exchange and more generally their scientific use.

Coding

If cataloguing is essential for the identification of the resources and for a fast comparison between them, coding is itself essential to the description of the interpretation of the content of the resource itself and also essential to its exploitation for specific studies.

If coding the transcription of video or audio material is indeed the clarification of an interpretation, then one notes here as well a great variety in the practices.

The inventories made during the EAGLES, MATE and ISLE successive initiatives have demonstrated the difficulty there is to apprehend the extent of the various coding systems. The specification of a universal software tool for annotation can be an access path to resources coded in a standardized way, as was suggested by ISLE.

But this does not make it any the less necessary for us to undertake a communal and consensual work to categorize, name and structure the phenomena found within the resources if we wish a true interoperability of the data with a view to multiple and future exploitations. Thus, we must now start standardizing the coding of oral corpora similarly to what has already been undertaken on written corpora with the TEL.

If we suppose that a compromise on the coding and cataloguing of oral data has been obtained, it is then necessary to give ourselves rules and tools to check the compliance of specific data with established principles. Therefore, in this symposium, we also wish to report on the practices used to control the quality of the resources.

Topics of interest / Thématiques

- | | |
|---|---|
| - description and cataloguing of oral resources | - description et référencement des données orales |
| - diffusion | - diffusion |
| - tools design and architecture | - spécification d'outils |
| - scientific exploitation of data | - exploitation scientifique |
| - archiving | - conservation |
| - corpus editing | - édition de corpus |
| - annotation | - annotation |
| - versionning | - versionning |
| - cataloguing and coding standards | - standards de catalogage et de codage |
| - ressources comparison | - comparaison des ressources |
| - multimodal and multimedia transcription | - transcription multimodale et multimédia |
| - annotation schemes | - schémas d'annotation |
| - interoperability | - interopérabilité |
| - evaluation, quality control | - évaluation, contrôle qualité |

Program Committee / Comité de programme

Jean-Yves Antoine (Université de 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, Paris)
Laurent Romary (MPI Berlin - INRIA)
Emmanuel Schang (Université d'Orléans)
Richard Walter (CNRS, Orléans)
Peter Wittenburg (Max-Planck-Institute for Psycholinguistics, Nijmegen)

Organisation Commitee / Comité d'organisation

Serge Heiden (ENS-LSH, Lyon)
Michel Jacobson (Ministère de la Culture, Paris)
Emmanuel Schang (Université d'Orléans)
Richard Walter (CNRS, Orléans)

REVUE DE SEMANTIQUE ET PRAGMATIQUE

RSP 21-22 (2007)

« La place du sens en linguistique »

Pierre Larrivé La source de la dimension qualitative des indéfinis de sélection arbitraire	9
Neiloufar Family Verbal Islands in Persian	23
Hélène Margerie Etude d'un adjectif qui se grammaticalise « grave »	49
Mirjana Miskavic-Lukovic Procedural semantics at work: the case of the English perception verbs <i>look, listen</i> and <i>see</i>	63
Daciada Vlad Du discours polémique à la polémique dans le discours	81
Pauline Merlet La Pragmatique Intégrée : application au discours lexicographique (le cas des articles <i>femme</i> et <i>homme</i> dans le <i>GR</i> et le <i>TLF</i>)	95
Veran Stanojevic & Tijana Asic One gift or two? The opposition distributive-collective and the syntax-semantic-pragmatic interface	105
Alfredo M. Lescano Vers une grammaire argumentative de <i>LE</i> et <i>UN</i>	119
Hugues Picavez Un exemple d'intégration sémantique de l'information pragmatique : les marqueurs modaux lexicaux	137
Bruno Courbon Usage(s) d'une approche pragmatique en sémantique diachronique : des lieux d'apparition et de stabilisation de l'innovation sémantique lexicale en français moderne	149
Nazaire Mbame Principes de l'adduction et abduction dans la catégorisation dynamique du langage.	175
Pascal Amsili L'annulation des implicatures et des présuppositions	193

RSP 23 (juin 2008)

Anne Theissen <i>Ledit N</i> : des emplois anaphoriques originaux	7
Laurence Kister & Evelyne Jacquey Traits sémantiques et anaphores pronominales	23
Patricia C. Hernández L'emploi locatif <i>sur</i> + ville	35
Pierre Cadiot & Abdelaali Talmenssour Dynamique sémantique du texte proverbial.	55
Vanda Mikšić Les actes de silence à la lumière de la théorie des actes de langage	75
Catalina Fuentes Rodríguez Invalidación, dialogismo y cortesía: Es un decir	95
Bénédicte Laurent Signifiante du nom de marque et de produit Ressorts sémantiques pour un effet pragmatique	121

RSP 24 (décembre 2008)
« Sémantique de la coordination »

Christian Bassac Sémantique de la coordination : une brève introduction	7
Frédéric Lambert & Lionel Clément Présentation	25
 <i>Sémantique des marqueurs de la coordination</i>	
Carlotta Viti The meanings of coordination in the early Indo-European languages	35
Katja Jasinskaja & Henk Zeevat Explaining Additive, Adversative and Contrast Marking in Russian and English	65
Anna Orlandini & Paolo Poccetti Liens de coordination : une approche sémantique à travers les langues anciennes	93
Frédéric Lambert Les intermittences du <i>ou</i> en français	115
 <i>Sémantique des éléments conjoints dans les structures coordonnées</i>	
Caterina Mauri The parallelisms of clausal coordination	145
Anne Abeillé & François Mouret Quelques contraintes sur les coordinations elliptiques en français	177
Mehmet Cicek & Christian Bassac Factorisation des affixes de pluriel et de possessif en turc.	207
Lionel Clément Zeugme sémantique	231

Numéro 25-26 (année 2009)

« Espace / Temps

(ou interprétations temporelles / interprétations spatiales ?) dans les langues »

Jean-Claude Anscombe

« Pour une autre classification des états »

Denis Le Pesant

« Modes d'action, diathèses et portée des ajouts temporels dans les verbes dits de déplacement »

Andrée Borillo

« La notion de "dynamique des forces" dans un emploi non-spatial de *sur* et *sous* en français »

Pierre Jalenques

« Valeurs spatiales et non spatiales dans le domaine des verbes »

Anne Le Draoulec & Denis Vigier

« *Dans* suivi d'un nom de partie de la journée : au croisement de l'espace et du temps »

François Nemo

« Interprétations temporelles et signification (non temporelle) : de *toujours* à *tout* »

Aude Rebotier

« Représenter le temps dans l'espace : les schémas pour la temporalité de l'énoncé »

Marcel Vuillaume

« *Maintenant*, un indexical pur et dur ? »

Samira Moukrim

« L'expression du 'présent actuel' en arabe (marocain), berbère (tamazight) et français parlés »

Adeline Patard

« L'imparfait dans le tour [(et) si *IMP?*] : entre temporalité et modalité ? »

Agnès Provot

« Le « futur dans le passé » en français et en allemand »

Inès Saddour

« The expression of progressivity in Tunisian Arabic: A study of progressive markers in oral retellings of simultaneous situations »

Pierre-Patrick Haillet

« Repérages temporels construits par le discours et approche polyphonique »

Mongi Kahloul

« *Après tout* : d'un *Après* temporel à un marqueur discursif de clôture »

Georgeta Cislaru & Michelle Lecolle

« Noms propres de lieux habités, espace et temporalité »

Numéro 27 (juin 2010)
« Des linguistes et des outils : au-delà des corpus »

Le Gallois & Leroy

« Présentation »

J. François

« Le repérage des segments répétés et la polysémie lexicale »

Ferrari & Legallois

« Modélisation d'un jeu de langage en vue d'explorations textuelles... »

Loiseau & Gréa

« Dictionnaires, théorie des graphes et structures lexicales »

Bertels, Speelman & Geeraerts

« La corrélation entre la spécificité et la sémantique dans un corpus spécialisé »

Recherches actuelles

Duguine & Irurtzun

« Opérateurs d'exclusivité dans les questions et réponses en basque »

Rudel & Mazaleyrat

« A propos d'un curieux adjectif : approche sémantico-cognitive de l'adjectif *curieux* »

Sonia Gomez & Jordana Ferary

« Énonciation proverbiale en français (XVI^e-XVII^e siècles) »

CORELA

Cognition, Représentation, Langage

Présentation de la revue en ligne et du site

CORELA (*Cognition, Représentation, Langage*) est la revue du *Cercle Linguistique du Centre et de l'Ouest* créée en 2000 et dirigée depuis par Gilles Col, Professeur à l'Université François-Rabelais. C'est une revue semestrielle et en accès libre, hébergée par la MSHS de l'Université de Poitiers.

L'une des originalités de CORELA est son caractère non thématique puisque y sont publiées des contributions sur des thèmes ouverts, qui abordent des questions de langue (grammaire, lexicque, syntaxe, phonologie, etc.), mais aussi de langage, de représentation, de modélisation, de sémantique. Dans la mesure du possible, sont regroupés des articles complémentaires par leur approche, leur méthodologie, leur corpus sans que soit imposée une thématique.

CORELA assure 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 PUR). Le choix du support électronique répond à la volonté de confronter rapidement des points de vue, d'accroître l'accessibilité et d'augmenter la périodicité des publications.

CORELA comporte 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, comptes rendus de manifestations scientifiques ou d'articles, pré-publications, textes de conférences, annonces de colloques, activités doctoriales...). Il n'y a pas de publications de thèses en l'état mais les auteurs sont encouragés à soumettre un article fondé sur leur recherche doctorale.

L'objectif de CORELA est de promouvoir les discussions scientifiques et théoriques les plus larges possibles. Le support électronique, qui facilite publication et diffusion, a permis d'en faire plus qu'une revue en ligne, une banque de connaissances et un lieu de confrontation des méthodes et des résultats.

CORELA a depuis le début associé étroitement des enseignants-chercheurs de l'Université François-Rabelais à ses activités, tant au niveau de la direction éditoriale (directeur : Gilles Col, co-directeur : Jean-Michel Fournier) que du travail d'évaluation des propositions d'articles ou de numéros spéciaux soumis aux différents comités (voir organigramme).

A ce jour, CORELA a publié 8 volumes de deux numéros chacun, ainsi que 10 numéros thématiques (dont trois issus de travaux ou de colloques tenus à l'Université de Tours, en 2005 et 2006). Le site de la revue a été visité par plus de 20.000 lecteurs en 2009. Elle est répertoriée par le portail du Directory of Open Acces Journals qui regroupe à l'Université de Lund plus de 5.000 revues en accès libre couvrant tous les domaines scientifiques.

Organigramme de la revue CORELA

Directeur de publication

Gilles Col (Université François-Rabelais, Tours).

Co-directeurs de publication

Jean-Michel Fournier (Université François-Rabelais, Tours)

Daniel Roulland (Université de Haute-Bretagne, Rennes)

Secrétaire éditorial

Franck Zumstein (Université Paris 7).

Comité international

Dirk Geeraerts (Université Catholique de Louvain, Belgique)

Simos Grammenidis (Université Aristote, Thessalonique, Grèce)

Hans Petter Helland (Université d'Oslo, Norvège)

Bert Peeters (Université Macquarie, Australie)

Len Talmy (Université d'Etat de New York, USA)

Comité scientifique

Mario Barra-Jover (Paris 8), André Borillo (Toulouse 2), Didier Bottineau (CNRS, Paris X-Nanterre), Philippe Bourdin (Toronto, Canada), Philipp Carr (Montpellier 3), Jean Chuquet (Poitiers), Gérard Deléchelle (Tours), Jean-Louis Duchet (Poitiers), Jacques François (Caen), Georges Kleiber (Strasbourg 2), Frédéric Lambert (Bordeaux 3), Nicole Le Querler (Caen), Jean-Jacques Lecercle (Paris X-Nanterre), Denis Maurel (Tours), Salah Mejri (Paris-Nord), Claude Muller (Bordeaux 3), Peter Roach (Reading, UK), Anne Trévisse (Paris X-Nanterre), Claude Vandeloise† (LSU, Bâton-Rouge, USA), Bernard Victorri (CNRS, ENS).

Comité de lecture

Jean Albrespit (Bordeaux 3), Nicolas Ballier (Paris-Nord), Hélène Chuquet (Poitiers), Catherine Collin (Nantes), Alain Delplanque (Tours), Hamida Demirdache (Nantes), Alain Deschamps (Paris 7), Nathalie Garric (Tours), Françoise Gayral (Paris-Nord), Thierry Grass (Strasbourg 2), Stéphane Gresset (Paris 7), Daniel Kayser (Paris-Nord), Jean-Charles Khalifa (Poitiers), Anne Lacheret-Dujour (Paris X-Nanterre), Florence Lefeuvre (Paris 3), Dominique Legallois (Caen), Michel Paillard (Poitiers), Gérard Mélis (Paris 7), Amina Mettouchi (EPHE), Véronique Rauline (Paris X-Nanterre), Stéphane Robert (CNRS, INALCO), Sylvester Osu (Tours).

CORELA

(Présentation des numéros)

2007, V, 1

Le premier numéro du volume 5 de CORELA s'ouvre par un article de Laurence Kister et Evelyne Jacquey qui proposent une manière d'acquérir des informations sémantiques pour résoudre les anaphores où l'expression référentielle est de la forme NdeN et l'expression anaphorique est un pronom relatif sujet. Le second article, de Viviane Arigne, aborde le problème de la modalité itérée dans la double perspective de la grammaticalisation et de la polysémie au travers de la description sémantique de divers emplois de l'auxiliaire de modalité *should* en anglais contemporain.

2007, V, 2

La revue CORELA fête son cinquième anniversaire avec la publication de ce nouveau numéro, et c'est encore une fois la diversité de la recherche en sciences du langage qui caractérise cette nouvelle publication puisque sémantique, traitement automatique des langues et grammaire sont représentés, parmi d'autres domaines de notre discipline. Vous pouvez effectivement lire un article de P.-A. Somé sur la polysémie et la recherche d'invariants du verbe « di » en dagara, un autre d'A. Dziadkiewicz qui présente un modèle théorique de la traduction automatique de phraséologismes pragmatiques, et enfin un article d'E. Roussel qui met en évidence certaines propriétés pragmatiques d'une marque aspectuelle de l'anglais (*past perfect*) qui font passer au second plan les propriétés habituellement attribuées à cette forme.

2008, VI, 1

L'objectif de CORELA est d'ouvrir la revue à tous les domaines de la linguistique. C'est ce que ce nouveau numéro tente de faire en publiant quatre articles qui touchent à des questions assez variées comme celles des locutions (article de Jeanne Aptekman sur « si »), des stratégies argumentatives (article de Sara Boutouhami et Daniel Kayser sur les descriptions d'accidents de la route), des structures argumentales (article de Malika Kaheraoui sur « faire + infinitif ») et enfin la question des prépositions (article de Laurence Vincent-Durroux sur l'utilisation des prépositions par de jeunes sourds oralistes).

2008, VI, 2

Outre deux articles qui abordent les phénomènes de discordance entre morphologie et syntaxe (Blandine Pennec) ou les difficultés communicationnelles posées par les féminisations des noms de métiers (Béatrice Fracchiolla), les autres articles de ce nouveau numéro de CORELA s'ouvrent à différents domaines du langage et la cognition. Ils portent effectivement sur la place du langage dans la cognition humaine, à travers l'étude de la compréhension d'énoncés arithmétiques et son impact sur la résolution des problèmes posés (Sylvie Akiguet-Bakong), et à travers la question plus générale de la place du langage dans la biologie humaine (Nicolas Bourguignon).

2009, VII, 1

C'est à différents aspects de la phonologie que ce nouveau numéro de CORELA est en grande partie consacré : l'article de Sylvester Osu et Nick Clemens traite de la question des nasales en ikwéré, celui de Riitta Välimaa-Blum présente la question des contraintes phonotactiques envisagées d'un point de vue cognitif, et enfin celui de Véronique Abasq étudie deux préfixes anglais (pre- et post-) dans une perspective morpho-phonologique. Les deux autres articles de ce numéro sont consacrés à la presse sous l'angle de la modélisation de la dynamique énonciative (Marie Chagnoux) et de la dénomination événementielle et du nom propre (Laura Calabrese Steimberg).

2009, VII, 2

Les deux études présentées dans ce second numéro du volume 7, portent respectivement sur les langues sourdes et l'émergence de nouveaux aspects linguistiques, tels que les commentaires métalinguistiques, les emplois métaphoriques, l'humour et les énoncés complexes (L. Vincent-Durroux) , et sur la relation entre style et identité à travers l'étude de la variation stylistique envisagée comme une stratégie verbale à part entière (L. Buson et J. Billiez).

2010, VIII, 1

CORELA publie dans son nouveau numéro deux articles qui portent sur la langue anglaise. Dans l'un, Mathilde Pinson montre que le GN « the way » est en passe de devenir une locution conjonctive à part entière ; dans l'autre, Mickaël Vallée analyse la structure PI for P2, et tente de déterminer si l'argumentation qui consiste à définir P2 comme une cause et / ou comme une proposition secondaire, voire optionnelle, est fondée.

CERLICO

Réseau des universités du *Cercle Linguistique du Centre et de l'Ouest*

Le CerLiCO est composé d'une douzaine d'universités et de laboratoires de recherche du Grand Ouest implantés à :

- Angers
- Bordeaux (ERSàB)
- Brest (CRBC)
- Caen (CRISCO)
- La Rochelle
- Le Mans
- Limoges (CERES)
- Lorient-Vannes
- Nantes (LING)
- Poitiers (FORELL)
- Rennes (ACE)
- Orléans (LLL)
- Tours (LLL)

Les objectifs du CerLiCo :

- développer, dans le domaine de la linguistique, la communication scientifique entre ses membres, à quelque école de pensée linguistique qu'ils appartiennent,
- faciliter l'intégration des jeunes chercheurs à la vie scientifique de la discipline,
- organiser des colloques et rencontres entre linguistes en collaboration avec les universités et centres de recherches, notamment dans le centre et l'ouest de la France,
- publier les actes de ces colloques et rencontres ainsi que d'autres travaux utiles à la recherche linguistique et à la diffusion de ses résultats,
- faire connaître au niveau international la recherche linguistique menée en France et susciter la participation de chercheurs étrangers aux activités de l'association.

Dans ce contexte de coopération et d'échanges, l'Université François-Rabelais participe activement depuis plus de vingt ans à la vie de l'association :

- en accueillant tous les ans l'assemblée générale d'automne,
- en ayant organisé trois colloques de l'association (1996, 2002, 2010), les deux premiers publiés aux PUR (Réseau des Universités de l'Ouest Atlantique dont font partie les Presses Universitaires François-Rabelais), le troisième (voir annexe 24^e colloque du CerLiCO) devant suivre,
- et en ayant participé à la fondation et à l'animation de la revue CORELA – revue du CerLiCO – depuis 2003.

La dynamique de fusion du LLL a conduit la partie orléanaise à rejoindre le Cercle à partir de 2008, en même temps que s'effectuait le rapprochement entre les deux universités.

Le 24^e colloque international du CerLiCO

Transcrire, Ecrire, Formaliser

Le 24^e colloque

Le colloque annuel du *Cercle Linguistique du Centre et de l'Ouest* s'est tenu à Tours les 4 et 5 juin 2010. Il a réuni une soixantaine de participants, de différents pays : France, Suisse, Grèce, Italie, Royaume-Uni, Japon. Le colloque a confirmé l'implantation tourangelle du CerLiCO grâce à l'engagement des enseignants-chercheurs et des doctorants du laboratoire, ainsi que des étudiants du master de sciences du langage « Linguistique avancée et applications », dans la préparation du colloque. Cet engagement s'ajoute à celui existant depuis de longues années, à tous les niveaux de l'association CerLiCO.

Le colloque accueillait un collègue de l'université Aristote (Thessalonique, Grèce), Simos Grammenidis, pour une conférence plénière sur la traduction : « Traduire : transcrire, ré-écrire ou ré-énoncer ? ». Sa conférence, ainsi qu'une sélection des communications présentées au colloque, seront publiées dans le Volume 24 des *Travaux Linguistiques du CerLiCO*, à paraître en juin 2011. Les Actes des colloques du CerLiCO paraissent depuis 1988 aux Presses Universitaires de Rennes (23 volumes parus à ce jour, dont le dernier en juin 2010) qui anime le Réseau des Universités de l'Ouest Atlantique dont font partie les Presses Universitaires François-Rabelais. Les articles issus des communications sélectionnées pour le colloque font chacun l'objet, avant publication, d'une double relecture.

Le thème traité cette année à Tours sera, conformément à la tradition du CerLiCO, le même que celui du prochain colloque de l'association qui se déroulera à Orléans en juin 2011, renforçant ainsi les activités communes des deux équipes du LLL. Outre le soutien financier du CerLiCO, de l'Université et du LLL, la région Centre s'est engagé et a soutenu notre manifestation.

Organisation du colloque

Président : Gilles Col (Tours)

Co-responsables : Jean-Michel Fournier (Tours), Sylvester Osu (Tours)

Comité scientifique :

Gilles Col, PR, Université François-Rabelais, Tours / EA 3850 LLL

Véronique Abasq, MC, Université François-Rabelais, Tours / EA 3850 LLL

Michel Achard, Professeur Assistant, Rice University, Houston, Texas, USA

Gabriel Bergounioux, PR, Université d'Orléans / EA 3850 LLL

Catherine Collin, MC, Université de Nantes / EA 3827 LING,

Jean-Louis Duchet, PR, Université de Poitiers / EA 3816 FoReLL,

Jean-Michel Fournier, PR, Université François-Rabelais, Tours / EA 3850 LLL,

Jacques François, PR, Université de Caen / EA 4255 CRISCO,

Simos Grammenidis, MC, Université Aristote, Thessalonique, Grèce,

Frédéric Lambert, PR, Université Bordeaux 3 / UMR 5610 ERSàB,

Sylvester Osu, MC, Université François-Rabelais, Tours / EA 3850 LLL,

Véronique Rauline, MC, Université Paris Ouest Nanterre / EA 370 CREA,

Noëlle Serpollet, MC, Université d'Orléans / EA 3850 LLL

Appel à communications

TRANSCRIRE ÉCRIRE FORMALISER

Pour se constituer comme science, la linguistique opère une rupture avec la philologie. En abandonnant l'édition et le commentaire des textes, elle établit son champ d'observation en dehors des écritures vernaculaires et compare des formes sonores. Elle ouvre ainsi un accès à toutes les langues (et à tous les états de langue), qu'elles aient été ou non consignées par écrit. Le geste fondateur de la discipline est présent aussi bien dans la reconstruction des comparatistes, qui proposent une notation de formes conjecturées, que dans le travail de terrain de Boas, Bloomfield et Sapir en Amérique du Nord, de Troubetzkoy dans le Caucase.

TRANSCRIRE, c'est interroger quelle signification revêt la conversion d'un signal sonore en une représentation graphique. Se trouvent concernés, en plus des questions de phonétique et de phonologie, les aspects figuratifs (sémiologie, langue des signes...), acoustiques (capture et traitement du son, visualisations...) et anthropologiques (introduction de l'écriture dans les civilisations à tradition orale, les graphies comme mode privilégié de la connaissance savante). Cette transcription recourt à des moyens variés (pictogrammes, syllabaires, alphabets...) dans une relation complexe avec la mise en forme des langues. Les idéogrammes imprègnent la culture chinoise dont ils assurent l'unité malgré la dialectalisation alors que la notation de l'arabe préserve la transparence de ses mécanismes apophoniques. L'écriture a des répercussions dans les représentations cognitives des auditeurs comme dans la structure sociale qui s'y trouvent inscrites en retour.

ECRIRE, c'est poser la question des relations entre un système conventionnel de représentation graphique et la structure (phonologique, morphologique, syntaxique, sémantique...) des langues. Les correspondances, phonie/graphie, les conventions orthographiques, jusqu'à la création de l'API se situent dans une notation analogique et descriptive qui a permis la production d'instruments tel que le dictionnaire sans expliquer le fonctionnement des langues. La linguistique s'affranchit de ces écritures reçues au moment où elle opère une distinction radicale entre ce qui s'entend et la structure des langues. Celle-ci n'est plus la consignation d'une observation mais le produit d'une analyse qui requiert à son tour une notation spécifique : une formalisation.

FORMALISER, c'est répondre à la nécessité d'un principe d'explication pour construire un système (et non plus un inventaire) ayant ses règles propres et qui vaut universellement dans le temps, dans l'espace et dans la société. Linguistique « algébrique » (Saussure) dont une partie des propriétés est justiciable d'un traitement automatique.

TRANSCRIRE, ECRIRE, FORMALISER constituera le programme des colloques du CerLiCO qui auront lieu à Tours en juin 2010 et à Orléans en juin 2011 afin d'appréhender l'objet de la linguistique au moment de sa constitution, dans sa genèse et ses effets.

(appel publié en juin 2009, sélection des communications en novembre)

Programme

Vendredi 4 juin 2010

9h00-9h15 **Accueil / discours**

9h45h – 10h25 : *Ecrire les nombres : quelles analyses pour la linguistique ?*

Sophie Saulnier, Modyco (CNRS / Université Paris 10-Nanterre)

10h25-11h05 : « *Formaliser* » *ACTUALLY: Reformulation, Formalisme, ou autre ?*

Lucie Gournay (Université Paris 12)

11h05 – 11h25 **pause**

11h25 – 12h05 : *Dans quelle langue transcrire les noms propres ?*

Sarah Leroy (Modyco CNRS / Université Paris 10-Nanterre)

14h00-14h40 : *Transcrire sans interpréter ? Illusion, nécessité ou possibilité*

Elisabeth Delais-Roussarie, H. Yoo (CNRS / UMR 7110-LLF, Université Paris-Diderot)
et B. Post (University of Cambridge, RCEAL (Research Centre for English and Applied
Linguistics))

14h40-15h20 : *Écrire une langue sans forme écrite: réflexions sur l'écriture et la transcription de la Langue des Signes Italienne (LIS)*

Claudia Savina Bianchini, Université de Paris 8, Saint-Denis, Università degli Studi di
Perugia, Istituto di Scienze e Tecnologia della Cognizione (ISTC), CNR, Roma

15h40-16h20 : *Productions langagières bilingues et bimodales d'enfants sourds : quels outils de transcription, pour quelles analyses ?*

Agnès Millet et Isabelle Estève (Université Stendhal Grenoble 3/ Lidilem)

16h20-17h : *'Textos sourds' en Métropole et à La Réunion : « Zones flottantes » du français ordinaire et spécificités du contexte de surdité*

Marion Blondel (UMR7023, CNRS-Paris 8), Jeanne Gonac'h, LIDIFRA, Université de
Rouen, Gudrun Ledegen, Université de St Denis de La Réunion, Julia Seeli, Université de
St Denis de La Réunion

17h00 **Assemblée générale de l'association CerLiCO**

Samedi 5 juin 2010

9h30-10h30 : *Traduire: transcrire, ré-écrire ou ré-énoncer?*

Conférence de Simos Grammenidis

Université Aristote de Thessaloniki

10h40-11h20 : *À la poursuite d'outils spécifiques, ou : comment transcrire l'acoustique ?*

Chantal Rittaud-Hutinet (Paris 3 / EA 1483)

11h20 – 12h00 : *Transcription de corpus d'apprenants multilingues de FLE et analyse interphonologique : enjeux méthodologiques*

Isabelle Racine, Genève, (Suisse), Sylvain Detey, Waseda University (Japon), SILS et Université de Rouen, Françoise Zay Université de Genève (Suisse) Yuji Kawaguchi, Tokyo University of Foreign Studies (Japon)

14h00-14h40 : *La controverse sur la transcription de l'anglais australien : une question identitaire*

Marjolaine Martin (Université François-Rabelais /LLL)

14h40-15h20 : *De la transcription à l'écriture : comparaison entre les cas du picard et du valaque*

Marijana Petrovic et Jean-Michel Eloy (Université de Picardie)

15h20-16h00 : *Quels systèmes de transcription phonétique et phonologique adopter dans le continuum dialectal italien ? Comment transcrire le polymorphisme dans les aires dialectales de l'Emilie occidentale (Italie du nord) ?*

Louis Begioni (Université Lille 3)

16h - 16h30 Clôture du colloque

Collaboration avec le laboratoire LaTTiCe

UMR 8094 (CNRS, ENS & Université Paris 3)

Un groupe de travail est animé au LaTTiCe par un chercheur du LLL, Gilles Col, autour de la modélisation de la construction dynamique du sens. Ce groupe réunit des chercheurs du CNRS essentiellement, Bernard Victorri (DR), Jeanne Aptekman (post doc), Stéphanie Girault (IR) et Thierry Poibeau (CR). L'objectif est de proposer une formalisation compositionnelle gestaltiste pour décrire la construction du sens au fur et à mesure du déroulement du discours.

Les recherches du LaTTiCe portent principalement sur la langue française dans une perspective synchronique, mais aussi diachronique et comparative interlangues. Les données linguistiques – incluant une large part d'études sur corpus – sont appréhendées dans une perspective de description, de modélisation et d'expérimentation. Ces recherches s'inscrivent dans les débats cognitifs et anthropologiques actuels sur les rapports du langage et de l'esprit.

L'orientation cognitive des travaux conduits en partenariat portent sur la diversité des formes linguistiques qui permettent d'exprimer de grandes catégories notionnelles telles que l'espace ou la comparaison. Ils concernent également la structuration des discours qui contribuent à l'ancrage des énoncés dans leur contexte et signalent les relations entre les situations dénotées par ces énoncés. Ces études sont menées dans le but (entre autres) d'explicitier et de modéliser les instructions interprétatives codées par ces expressions et de mesurer / expliquer la façon dont les sujets les traitent dans la compréhension en temps réel. C'est cette partie des recherches conduites au LaTTiCe qui est au centre de la collaboration avec le LLL.

Le groupe de travail s'attache effectivement à mettre en évidence les instructions fournies par les unités linguistiques d'un énoncé lors de leur perception dans le déroulement du discours. Ces instructions dynamiques sont formulées par leurs formes schématiques qui reposent sur le principe de convocation / évocation. Elles répondent à la double question : qu'est-ce qu'une unité *convoque* sur la scène verbale pour pouvoir construire du sens et qu'est-ce que cette unité apporte à la construction de la scène – ce qu'elle *évoque* – en fonction des éléments qu'elle a convoqués et sur lesquels elle agit ? La finalité est de proposer une modélisation de la construction du sens qui donne une place centrale à la polysémie (Victorri & Col 2007, Col 2008a, 2008b, Col, Aptekman, Girault & Victorri, 2010, Aptekman & Co, 2010).

Un projet d'ouvrage consacré à l'exposé des principaux résultats ainsi qu'à la présentation d'une « grammaire instructionnelle » est actuellement en cours. La collaboration avec le LaTTiCe se poursuivra lors du prochain quadriennal (cf. le projet de l'axe de recherche SEC, « Sémantique énonciative et cognition ») en s'élargissant à d'autres chercheurs des deux laboratoires.

THEMATIQUE TRANSVERSALE « CONSTRUCTION DU SENS »

Rapport développé

Responsable : Sylvester N. Osu

Participants : V. Abasq, C. Agafonov, A. Cambourian, G. Col, G. Deléchelle, A. Delplanque, J.-M. Fournier, T. Gallèpe, N. Garric, D. Maurel, S. N. Osu, A. Piel, J.-P. Régis, F. Toupin, J. Constanza, P. Fournier, M. Martin, A. Ndione, E. Lecuit, J. M. Nagy, G. Potjans, I. Girard, S. Vanhoutte, L. Perbost, J. Longhi, E. Sauquet.

Participation des étudiants de M2.

Autres participations : Feifei Chen (sinophone et sinologue. Université de Franche Comté, Besançon), Samira Moukrim (LLL-Orléans).

1. Objectifs

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 de LLL Tours de se constituer un domaine d'action et d'interaction scientifique commun, à la fois donc lieu d'échange et de convergence, mais aussi ambition d'explorer les potentialités d'une linguistique plus intégrative.

2. Problématique

La plupart des recherches actuelles en linguistique intègrent d'une manière ou d'une autre, la dimension sémantique du langage. C'est le cas en particulier de la grammaire cognitive dont les représentants les plus connus sont R. Langacker, L. Talmy, C. Vandeloise, A. Wierbiczka ; de la Functional linguistics avec M.A.K Halliday, E. Sweetser, Bybee, K. Lambrecht, T. Givon ; de la Formal semantics à la Montague, de la Pragmatique avec O. Ducrot, J. Anscombre ; de l'Analyse du discours avec P. Charaudeau ; de la Sémantique générale de B. Pottier, de la Generative semantics de G. Lakoff, Postal et Ross ; de la Functional syntax à la Kuno, puis de la Dynamic syntax de R. Kempson et ses collègues de King's College London, pour ne citer que celles-là. Le point convergent de toutes ces approches, c'est que chacune tente 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. Toutefois, cette thématique centre son investigation non pas sur le résultat, autrement dit le produit, mais sur ce qui 'engendre' ce produit. Ainsi, elle s'inscrit dans l'optique théorique de la linguistique de l'énonciation telle qu'elle a été formulée et élaborée par A. Culioli et son équipe (cf. Culioli 1990, 1999 et ailleurs). En effet, non seulement la construction du sens se trouve au centre même des préoccupations de la linguistique de l'énonciation mais aussi cette approche

offre un outil théorique rigoureux pour la description fine, la caractérisation et l'explication des phénomènes. De plus, dans le cadre de la linguistique de l'énonciation, l'objet 'sens' est considéré comme le résultat de l'interaction de diverses opérations linguistiques, lesquelles sont représentées en surface par des formes appelées également des marqueurs (au sens de marqueurs d'opérations). Ces formes ou marqueurs peuvent être la prosodie, les mots, les syntagmes, les locutions ou les propositions dans des énoncés. Ces représentations en surface sont effectuées par le sujet énonciateur relativement à la situation d'énonciation. Toute recherche portant sur la construction du sens se doit désormais de prendre ces deux derniers paramètres en considération.

3. Diversité des langues

La recherche au sein de cette thématique transversale est abordée à travers plusieurs langues qui y sont représentées à savoir, l'allemand (langue germanique de la branche indo-européenne), l'anglais (langue germanique de la branche indo-européenne), le dagara (langue gur de la branche Niger-Congo), l'espagnol (langue latine de la branche indo-européenne), le français (langue latine de la branche indo-européenne), l'ikwere (langue igboïde de la branche Niger-Congo) et le russe (langue slave de la branche indo-européenne). De ce fait, le travail qui y est mené se veut contrastif et en même temps, s'inscrit dans une optique typologique. Il est à rappeler que la linguistique de l'énonciation vise à appréhender le langage à travers la diversité des langues naturelles. Ce qui signifie que c'est en allant jusqu'au bout de la singularité des langues qu'on arrivera à cerner les mécanismes généralisables de l'activité langagière. Il a semblé opportun de tenter sur la base de cette diversité des langues, de cerner le mécanisme de la construction du sens à travers des phénomènes précis. Bref, il s'agit de rendre compte de la façon la plus précise possible, des manières spécifiques par lesquelles chacune des langues représente des phénomènes particuliers en rapport avec la construction du sens.

4. Réalisations

L'équipe a abordé dans les cinq années écoulées deux thèmes majeurs à savoir la dénomination (2005-2007) et la temporalité (2008-2010). Un résumé de notre investigation sur chacun des deux thèmes est présenté ci-dessous.

4.1 De la dénomination à la construction d'identité

Pour la période 2005-2007, l'équipe a choisi de travailler sur la problématique de la dénomination. D'abord, cette problématique nous semblait, à défaut d'être un prolongement, être en lien avec les travaux précédents sur les Noms propres par certains membres de l'équipe, et sur les verbes d'apparence par certains autres membres de l'équipe. Ensuite et surtout, la dénomination apparaissait comme un lieu où se 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 (voir en particulier, la prononciation des noms propres dans les différentes langues), la syntaxe et l'énonciation (processus d'identification, la négociation du sens). Le postulat était que seule la prise en compte de cette interaction permettrait de mieux rendre compte des phénomènes observés et ainsi définir plus précisément la dénomination.

La première phase d'exploration nous a permis de constater que la dénomination renvoie à la catégorisation des êtres et/ou des objets en ce sens que catégoriser, c'est identifier un élément à un groupe tout en marquant sa singularité. Cela revient à marquer une différence entre des objets ou des personnes, et donc, à construire une identité. Dès lors, la dénomination semble

construire l'identité à travers un processus de différenciation et inversement, marquer la différence à travers la construction d'une identité. Ce renvoi à la catégorisation nous a fourni un moyen de rapprocher la dénomination de la réduplication (voir par exemple en français, « je veux un café café ») tandis que la mise en jeu de la différenciation dans son fonctionnement rappelle les verbes d'apparence (cf. numéro spécial du *CORELA* coordonné par Delplanque, A. 2006). Par la suite, notre investigation a porté en parallèle sur la dénomination et sur la réduplication comme l'illustrent les calendriers de travail suivants :

2005-2006

19 janvier 2006: « Autour de l'identification » par G. Deléchelle

9 février : « Réduplication : état des lieux » par S. Osu

9 mars : « Classification et tabou » par J.P. Régis

13 avril : « Réduplication en Dagara » par A. Delplanque

11 mai : « Noms propres et expressions définies en tant que supports de la notion de personne » par T. Grass

8 juin : « Réduplication en russe » par C. Agafonov

2006-2007

12 octobre 2006: « Traitement phonologique des noms propres anglais » J.M. Fournier

9 novembre : « Constitution et exploitation d'un dictionnaire relationnel multilingue des noms propres » D. Maurel

7 décembre : « La représentation syntaxique des prépositions faibles du berbère de Chemini » M. Haiden

18 janvier 2007 : « Dérivés adjectivaux en -ish des noms d'animaux » Toupin & Lagorce

15 février : « Terminologie juridique : modèle bilingue canadien » T. Grass

15 mars : « Un tsunami médiatique et linguistique : de la dénomination à l'événement discursif » N. Garric

12 avril : « Réduplication en ikwere » S. Osu

10 mai : « Alternance morphologique systémique dans les noms propres en espagnol » A. Piel

21 juin : « Réduplication en chinois » Fei Fei Chen (Invitée, Université de Franche-Comté, Besançon)

Un peu mieux éclairée par les premiers travaux sur les deux phénomènes, l'équipe décida de reformuler son thème de recherche. De la dénomination l'on est passé à l'identité. Il s'agissait concrètement de cerner la part de la (L)angue dans la construction de l'identité d'une personne ou d'un objet. En d'autres termes, quels sont les procédés linguistiques qui rendent possible une telle construction d'identité. Mais la question s'est vite révélée très complexe aussi bien par les diverses acceptions (sens commun et sens technique confondus) qu'on attribue au terme d'identité, par sa diversité d'emplois (identité religieuse, identité politique, papier d'identité, etc.) que par la diversité de procédés qu'il recouvre. Nous l'avons tout de même abordée à travers les procédés de la dénomination et la réduplication dans les langues représentées dans notre équipe.

4.2 Bilan

Nous avons cherché à préciser les mécanismes spécifiques mis en jeu par chacun des phénomènes en prenant en compte des propriétés relevant du domaine lexical, de la phonologie, de la morphologie, de la syntaxe et de l'énonciation.

4.2.1 Dénomination

La dénomination a été abordée sous deux angles. Une première approche concerne le traitement automatique et le traitement phonologique des noms propres. Dans le premier cas de figure,

prenant comme exemple un dictionnaire relationnel multilingue, on s'est interrogé entre autres choses, sur les relations que les noms propres aussi bien dans une langue donnée que dans le passage d'une langue à une autre (la traduction), peuvent entretenir entre eux. Il s'agissait alors, de tenter de répondre à la question de la variation de noms que l'on peut rencontrer dans des textes lors d'une traduction. Par exemple, qu'est-ce qui dans un texte, permet de lier Mozart à Créateur de la flûte enchantée ? Ce type de travail exige que l'on précise les notions telles que synonymie, métonymie, hypéronymie. Or les relations entre noms propres permettent d'illustrer un système de catégorisation. Quant au traitement phonologique des noms propres, il s'agissait de montrer que l'accent final en anglais par exemple, n'existe que dans des mots provenant historiquement du français. La conséquence immédiate, c'est que l'identité linguistique des noms avec un accent terminal ou final en anglais peut facilement être établie. Il est à rappeler qu'une telle analyse a nécessité la connaissance de l'histoire du contact entre les deux langues. Une seconde approche se concentre sur les conséquences de l'acte de re-nommer. Or, re-nommer, cela consiste à balayer une unité pour la remplacer par une autre. Ainsi, on touche à la valeur identitaire de la personne, chose ou événement re-nommé. C'est ainsi par exemple que tsunami utilisé à la place de raz de marée à la suite de l'événement du 26 décembre 2004, puis étendu à d'autres domaines de la vie dans les discours médiatiques permet aux journalistes qui l'emploient de désigner un événement hors de propos. On parle désormais de tsunami politique pour signifier que l'événement politique est d'une grande ampleur, tsunami d'informations pour signifier que l'on est envahi par les informations et tsunami médical pour parler d'un problème médical qui fait trembler les pouvoirs en place.

4.2.2 *Réduplication*

A l'occasion de nos travaux sur la réduplication, nous avons mis en évidence le fait que ce procédé peut être distingué formellement et sémantiquement de la répétition ; que poser une forme (la base) et la reprendre (la copie) soit partiellement soit entièrement revient à introduire ce que dans la cadre de la TOPE, l'on appelle occurrence quelconque d'un domaine notionnel (base) et occurrence type d'un domaine notionnel (copie) ; puis que la juxtaposition des deux formes se traduit en termes d'identification. Ainsi, l'occurrence quelconque est identifiable à l'occurrence type. C'est ce mécanisme qui engendre non seulement une interprétation de type 'haut degré', 'le vrai', 'le parangon' mais aussi permet de construire une sous catégorie d'un domaine de référence. De ce fait, « un parfum pour les femmes femmes » (un spot publicitaire), c'est un parfum qui une fois acheté par une femme quelconque, affecte cette femme à une sous catégorie de femmes *par excellence*. On comprendra qu'un tel spot publicitaire n'a pour but que de pousser les femmes à acheter ledit parfum afin d'appartenir à la classe réservée aux femmes *par excellence*. Il a été également mis en évidence dans certaines langues que la réduplication à travers le processus d'identification d'une occurrence quelconque à une occurrence type, construit une propriété définitoire ou caractéristique d'un être humain ou d'un objet. Ainsi, lorsqu'en ikwere (langue du Nigeria), l'animal Ekeni dit que la nature lui donne « œil œil », il tend à dire que sa propriété définitoire se ramène à œil.

4.2.3 *Colloque international*

Encouragés désormais par nos observations, nous avons voulu voir s'il existait d'autres procédés dans des langues non représentées au sein de notre équipe et/ou comment et en quoi d'autres approches méthodologiques et théoriques permettraient d'aborder cette problématique que nous avons faite nôtre. Cela a débouché sur un colloque international intitulé « Construction d'identité et processus d'identification » et organisé à l'Université François Rabelais, Tours les 29 & 30 novembre, 2007. Organisé en quatre sessions parallèles, ce colloque a accueilli 66 communications orales plus 2 conférences plénières. Les diverses communications se sont

portées sur 12 langues à savoir, anglais, arabe marocain, arabe jordanien, espagnol, ewe, français, français de Côte d'Ivoire, ikwere, japonais, khmer, roumain et russe. Mais les communications se sont déroulées soit en français, soit en anglais. Quant au profil des participants, il était assez hétérogène : Américains, Canadiens, Nigériens, Belges, Italiens, Sénégalais, Espagnols, Français, Marocains, Japonais, Russes, Ivoiriens, Argentins et Jordaniens.

Ce colloque a révélé d'autres procédés par lesquels les langues du monde marquent et construisent l'identité. Ainsi, à côté de la dénomination et la reduplication, il y a aussi entre autres, l'hypallage, la paradiastole, l'antanaclase, le genre, la structure X entre les X, X chez les X, la structure X comme/tel que Y, le stéréotype, rupture de la chaîne anaphorique...

4.2.4 *Archivages*

Les deux conférences plénières ont été filmés par le service audio-visuel de l'Université François Rabelais, Tours et archivées sur le site de l'université :

http://lettres.univ-tours.fr/1190190229151/0/fiche___actualite/&RH=1192087652523

4.2.5 *Ouvrage*

Vu l'importance du résultat de ce colloque, nous avons décidé de recueillir la version rédigée et anonyme des communications présentées. Et après l'avis d'un comité de relecture que nous avons sollicité et mis en place, nous avons retenu 39 contributions, soit celles qui ont reçu un avis favorable de la part d'au moins de relecteurs. Revues et corrigées, ces contributions sont réunies dans un ouvrage de 623 pages, sous le titre « Construction d'identité et processus d'identification », publié chez Peter Lang (Berne, Suisse) en 2010. Les deux dernières pages de cet ouvrage présente la liste des relecteurs.

4.3 *Temporalité*

Pour la période 2008-2010, l'équipe a choisi de travailler sur la temporalité. Le programme comportait deux volets : 1) comment les langues et discours permettent-ils de construire le temps ? Dans ce cas, la temporalité est envisagée comme objet de description linguistique ; Concrètement, il s'agit de cerner les unités dans les langues qui marquent la temporalité. 2) comment le facteur temps est-il utilisé dans la description linguistique ? La temporalité est alors vue comme moyen de description linguistique. En effet, la temporalité a été largement abordée relativement à certaines langues de l'aire linguistique européenne mais il n'en est pas de même des langues relativement peu documentées et décrites telles que le dagara et l'ikwere. Nous voulions donc à travers une approche contrastive, faire une mise au point sur cette notion en linguistique. Les exposés pouvaient se situer dans l'un ou l'autre volets. L'exploration des concepts de forme schématique, de repérage et de situation d'énonciation, concepts au cœur de la TOPE, s'est avérée très utile à la réflexion. La puissance explicative de ces concepts tient précisément au fait qu'ils font de la variation un facteur indissociable de l'invariant et de ce fait, ils permettent un traitement unitaire et d'emblée syntaxico-sémantique des phénomènes. Ils permettent également de restituer la complexité de la temporalité souvent simplifiée à travers une représentation en axe. Les temps forts de notre réflexion ont été illustrés par quelques interventions. G. Col propose une analyse des prépositions « by » et « until » en anglais, soutient que ces éléments sont loin d'avoir un « sens premier » qui serait spatial ou temporel, mais qu'ils ont surtout un rôle instructionnel dans la construction du sens d'un énoncé, un rôle qu'il analyse comme leur forme schématique. A. Delplanque s'interroge sur la construction de la temporalité dans la conjugaison verbale en anglais comme en dagara, et S. Moukrim, dans une perspective contrastive, s'interroge sur la forme du présent en arabe marocain, berbère tamazight et français.

Mais si le fil conducteur de notre réflexion reste le rôle de la temporalité dans la construction du sens, certains exposés, en conformité avec les objectifs de l'équipe, ont toutefois été l'occasion de revisiter des travaux fondamentaux relatifs à la temporalité (c'est le cas, entre autres, du temps opératif chez Guillaume présenté par V. Mariscal et P. Bunte) alors que d'autres ont revisité des travaux antérieurs (c'est le cas de D. Maurel sur l'interprétation du temps verbal dans le TAL, puis sur la reconnaissance automatique des dates). Certains exposés se sont également demandés ce qui relève de la temporalité dans la grammaticalisation des unités (F. Toupin) tandis que d'autres ont voulu comprendre ce qu'il y a dans le préfixe « pre- » en anglais par exemple, qui peut permettre de renvoyer au locatif et au temporel (V. Abasq); mais aussi, comment le temps travaille la forme et le sens (J. Longhi) et enfin, le temps en tant que durée, c'est-à-dire longueur vocalique (M. Martin).

Sans prétendre avoir répondu à toutes nos questions, les divers exposés ont été en tout cas, l'occasion d'affiner nos questions et hypothèses, de porter un regard différent sur la temporalité, et ainsi renforcer notre objectif de faire collaborer à l'analyse de la construction du sens des spécialistes de disciplines linguistiques très différentes.

4.3.1 *Journée d'étude sur la temporalité*

En rapport avec ce programme, a été organisée à Tours le 5 mars 2010, une journée d'étude qui avait pour intitulé Journée d'étude Temporalité, sous la responsabilité de S. Osu et F. Toupin. Nous avons voulu à travers cette journée souligner à la fois la perspective typologique de nos travaux et la pluralité des approches comme l'illustre le programme ci-joint.

4.3.2 *Archivage*

L'ensemble des communications de la journée a été enregistré par le service audio-visuel de l'Université François Rabelais, Tours, et archivé sur le site de l'université, accessible via : http://lettres.univ-tours.fr/1266596192086/0/fiche___actualite/&RH=1179214274997

Journée d'étude sur la temporalité

Tours, 2009

Daniel Lebaud

Université de Franche Comté, LASELDI

Etude sur le français

« Les formes verbales du français : régimes énonciatifs et mises en scènes énonciatives des procès »

J'exposerai une recherche en cours, située dans le cadre de la Théorie des Opérations Prédicatives et Énonciatives, qui vise à articuler modes de repérages (régimes énonciatifs : "identité"(=), "différence"(≠), "rupture"(ω) et "ni identité ni différence ou identique ou différent"(*)) et modes de représentations des procès (mises en scène énonciatives : nécessaire, possible, contingent et intentionnalité, visée, Il y a). Il s'agit en conséquence, à partir d'une réflexion sur la distribution des formes verbales dans différents genres de textes (modes d'emploi, articles de presse, biographies, ...), d'une tentative de généralisation aux formes verbales, en rapport avec une réorganisation systémique, des notions subsumées sous l'appellation mises en scène énonciatives.

Stéphane Robert

LLACAN (UMR 8135, CNRS-INALCO)

Etude sur le wolof

« Le temps dans les langues africaines : l'exemple du système verbal du wolof »

Du point de vue grammatical, les langues africaines présentent des systèmes temporels assez divers que l'on ne peut ramener à un modèle unique. On présentera d'abord quelques caractéristiques remarquables relevées dans les systèmes temporels de ces langues : indications métriques, auxiliaires à valeurs aspecto-temporelles variées, amalgames entre temps et négation, entre temps et focus, orientation variable de la flèche du temps... On s'attachera ensuite à décrire plus précisément le système verbal d'une langue africaine parlée au Sénégal : le wolof. On montrera notamment que, dans cette langue, les indications aspecto-temporelles sont amalgamées à d'autres déterminations : il en résulte un système assez complexe à deux niveaux d'organisation et dans lequel la valeur d'accompli est construite de trois manières différentes. On soulignera également le rôle des types de procès dans la construction de la valeur temporelle des conjugaisons, ainsi que dans les glissements de valeur de certaines formes, du temps à la modalité.

Lotfi Abouda & Samira Moukrim

LLL (Université d'Orléans, EA 3850)

Etude sur le français

« Deux groupes de formes temporelles en français »

La morphologie du conditionnel, en tant que temps de l'indicatif, suggère l'existence dans ce mode de deux paradigmes de morphèmes « temporels » susceptibles de se concaténer. Une lecture possible de ces données suggère que la catégorie « temps », telle qu'admise habituellement, serait une notion générique synthétisant deux catégories dont la distinction implique la bipartition des temps de l'indicatif en deux sous-groupes distincts, l'un centré sur le présent, le second sur l'imparfait.

Silvio Liuzzi

CELIA (CNRS)

Etude sur le guarani

« Le temps en guarani »

Cette communication présentera le guarani dans la famille linguistique Tupi-guarani à l'époque pré-colombienne, coloniale, post-coloniale et actuelle ; une liste lexicale d'expressions temporelles ; les rapports temps / aspect / mode.

Laurent Gosselin

LIDIFRA (EA 4305) Université de Rouen

Etude sur le français

« De l'itération à l'aspect itératif »

Cette communication portera sur la nature de l'aspect itératif. Partant d'un examen critique de l'analogie, présentée dans quelques travaux récents, entre la quantification dans le domaine nominal et l'analyse de l'itération comme quantification plurielle sur les événements, j'examinerai la spécificité des propriétés aspectuo-temporelles de l'itération

Gary German

Université de Brest

Etude sur le gallois

« La temporalité et le système verbal du gallois : perspective diachronique »

Cette communication retracera l'évolution diachronique du système verbal du gallois depuis la période du moyen-gallois en soulignant notamment le passage d'un système synthétique à un système analytique, le développement de constructions périphrastiques liées (le plus souvent) à la progressivité, ainsi que les formes « bod » en B et en S, et, plus généralement, les conséquences que ceci a pu avoir sur la perception de la temporalité dans la langue moderne.

VARILING

La réponse à l'appel ANR Corpus 2006 du LLL et de ses partenaires

LE RÉSUMÉ

TRAITEMENT DES VARIATIONS LINGUISTIQUES DANS LES CORPUS (VARILING)

Le projet VARILING répond à une demande récurrente en linguistique : accroître une expertise sur les langues - en l'occurrence le français – qui intègre les variations et en permette l'analyse, qu'il s'agisse de la reconnaissance des données, de leur étiquetage ou des différents traitements dont elles peuvent faire l'objet. Les technologies ont été développées principalement à partir de corpus écrits et pour que de nouvelles connaissances, applicables à l'oral, puissent être élaborées, la communauté scientifique s'est souciée d'établir des recueils de parole circulante constitués en ressource.

Le plus important témoignage sur le français parlé avant 1980 est l'Enquête Socio-Linguistique à Orléans (ESLO 1), une collection de 200 interviews référencées (caractérisation sociologique des témoins, identification de l'enquêteur, date et lieu de passation de l'entretien), complétées par des captures sonores en contexte professionnel ou privé, soit au total 300 heures de parole et un corpus estimé à 4 500 000 mots. Une nouvelle enquête, ESLO 2, a été mise en chantier par le CORAL (Centre Orléanais de Recherche en Anthropologie et Linguistique, EA 3850) afin de constituer, à quarante années de distance, un corpus comparable dans les modalités de collecte et dans l'archive. L'objectif a été fixé à 400 heures environ de documents sonores, soit 6 000 000 de mots environ.

ESLO 2 a pour caractéristique d'être le seul programme en France (il en existe un équivalent au Québec) qui se propose de raisonner la dynamique sociolinguistique du français en bénéficiant d'un horizon de rétrospection de quatre décennies. L'étude est à même d'intégrer la variation diachronique et sociale et leur interaction ; elle est ouverte à une comparaison dialectale avec l'enquête de Lormont conduite par le CELITH-MODYCO.

Dès sa conception, ESLO 2 a été conçue pour préfigurer une référence dans un domaine qui, à l'échelle internationale, est en structuration et où l'adoption d'un format standardisé de *collecte*, de *conservation*, de *traitement* et d'*analyse* est confrontée à la multiplicité des développements et des normes. La certification est construite d'abord en tenant compte des pratiques en usage concernant la fabrication des corpus oraux en linguistique. A partir de l'échange entre les acteurs de la recherche, une synthèse des recommandations sera effectuée par le consensus des usagers et des experts, déterminant une conception du traitement d'ESLO 1 et de fabrication d'ESLO 2 qui permette d'exemplariser ces deux fonds pour de futurs corpus (et leurs traitements).

Si ESLO 2 a une visée cumulative (accroître la quantité de données pour assurer des comparaisons avec d'autres), l'enquête est aussi réflexive (accompagner l'enquête, le traitement et l'exploitation d'une analyse de l'expérience pour contribuer à la définition des normes). Cette conception concerne :

- une prospective sur l'exhaustivité des usages avec un calcul de représentativité,
- un inventaire des techniques de collecte (formats d'enregistrement et numérisation),

- une politique de formation des enquêteurs et d'information des témoins afin d'intégrer dans les critères de variation celle liée à l'enquêteur avec pour projet l'organisation d'une école d'été sur ce thème,
- un recueil des données concurremment à l'enrichissement en méta-données,
- un codage et un catalogage anticipant les principales requêtes émergeant en linguistique, mais aussi en sociologie, en anthropologie, en histoire, en info-com...
- une transcription avec alignement dans une perspective de normalisation,
- un étiquetage, avec catégorisation et lemmatisation (en particulier, recherches sur les problèmes rémanents de disfluences de l'oral et de co-référence anaphorique en situation de parole spontanée),
- une procédure d'anonymisation (l'identification des questions posées par l'anonymisation permettra la confection d'un vade mecum des éléments à prendre en considération à cette étape du travail à partir d'une recherche sur la détection des entités dénommantes (et pas seulement des entités nommées)),
- un stockage, avec archivage et indexation,
- une procédure de mise à disposition : la construction et la maintenance du site doivent assurer une libre consultation sur Internet (avec une convivialité et une ergonomie des applications, si possible dans une version multilingue),
- des données partagées : interopérabilité et protections, en liaison avec les propositions formulées dans le cadre du programme pour le catalogage et codage des corpus CatCod qui prolonge le travail de l'EPML50, « Corpus d'interaction langagière ». Les spécifications retenues seront transmises au consortium « Text Encoding Initiative » (TEI) à titre de proposition.

Au-delà, seront dessinées la mise en place du suivi (maintenance, jouvence et sécurité) et les applications. Il ne s'agit pas d'anticiper les analyses, mais de les rendre possibles.

Ce projet témoigne d'une ambition forte. Constituer en prototype, à toutes les étapes de sa réalisation, un corpus qui puisse se situer au même niveau, qualitatif et quantitatif, y compris par sa dimension patrimoniale, que les grands corpus oraux fabriqués, ou en cours de fabrication, en Europe et dans le monde. En trois ans, l'exécution de ce projet doit contribuer de façon décisive à la structuration de la communauté, en partenariat avec l'ensemble des laboratoires, dans une démarche associant l'accumulation des connaissances et la critique de leur constitution, l'analyse des outils et la disponibilité des données, le dialogue des chercheurs et l'échange de corpus devenus interopérables autour d'une thématique : la prise en compte des variations dans le langage.

L'EXPOSÉ DES MOTIFS

1. Présentation des corpus

On se propose d'exploiter et constituer des corpus qui permettent de mettre en évidence, en jouant sur la comparabilité des données, les configurations de la variation et les solutions de traitement. Pour ce faire, on procédera de quatre façons, concurremment et complémentaires :

- en traitant un corpus ancien de français oral à « récupérer » (ESLO 1) par différence avec un corpus de français contemporain à constituer suivant les standards en cours d'élaboration, auxquels il est prévu qu'ils servent de banc d'essai (ESLO 2) ;
- en éprouvant ces deux corpus pour tester l'alignement et le traitement parallèle de données commensurables quoique hétérogènes ;
- en rapprochant les corpus ESLO 1 et ESLO 2 élaborés à Orléans d'autres corpus du français (Lormont / CELITH ; Phonologie du Français Contemporain) ;
- en confrontant ces corpus de français à des corpus de langues non écrites (corpus LACITO, créoles à base portugaise) et d'Amérique (langues de Guyane).

1.1 *ESLO 1* : un travail pionnier

L'Enquête Socio-Linguistique à Orléans (désormais : ESLO 1) a été conduite en 1968 par des universitaires britanniques avec une visée didactique : l'enseignement du français langue étrangère dans le système public d'éducation anglais. Elle comprend environ 200 interviews, toutes référencées (caractérisation sociologique des témoins, identification de l'enquêteur, date et lieu de passation de l'entretien), soit au total plus de 300 heures de parole incluant pour moitié des interviews en face à face et pour moitié une gamme d'enregistrements variés (conversations téléphoniques, réunions publiques, transactions commerciales, repas de famille, entretiens médico-pédagogiques, etc.).

ESLO 1 couvre l'ensemble des catégories socio-professionnelles, hommes et femmes, avec plusieurs locuteurs originaires de différentes régions. C'est un échantillon des formats de la communication, des tâches linguistiques, des types de discours selon une approche essentiellement dialogique. Ce corpus représente, par son ampleur, sa rigueur et sa cohérence, le plus important témoignage disponible sur le français parlé avant 1980 (corpus estimé à 4 500 000 mots). Si les fins de sa constitution étaient linguistiques, ESLO 1 est un témoignage unique sur les jugements concernant mai 68 vu de la province ou sur les représentations collectives de la cité à cette époque.

Dans sa version première, transférée d'Angleterre il y a une quinzaine d'années et déposée aux Archives du Loiret, ESLO 1 réunit les bandes magnétiques, un catalogue dactylographié, quelques centaines de feuillets de transcription manuscrites (d'une qualité inégale) et les fiches d'identification des locuteurs. Actuellement, les documents sonores ont été recolligés et complétés (la conservation avait été défectueuse), numérisés à partir des enregistrements et une indexation et un premier catalogue informatisé a pu être réalisé. Parallèlement, l'exploitation exhaustive d'un sous-ensemble a été entreprise au point de rencontre de données linguistiques variationnistes et cognitives (description d'une tâche). Il reste à transcrire et baliser l'intégralité du corpus selon des principes ayant valeur de normalisation afin de mettre l'ensemble des données à la disposition de la communauté scientifique dans un format qui en permette une exploitation fiable, optimale et intensive, y compris pour des applications industrielles après sélection des contenus (en particulier concernant l'anonymisation).

1.2 *ESLO 2* : construire de la comparaison

En partant des acquis d'ESLO 1, une nouvelle enquête, dénommée ESLO 2, a été mise en chantier par le CORAL (Centre Orléanais de Recherche en Anthropologie et Linguistique, EA 3850). Il s'agit, à quarante années de distance, de constituer un corpus comparable dans le produit attendu et dans les modalités de la collecte : l'objectif a été fixé à 400 heures environ de documents sonores qui totaliseraient approximativement 6 000 000 de mots. Réunis, ESLO 1 et ESLO 2 formeront une

collection de 700 heures d'enregistrement, soit plus de 10 000 000 de mots, ce qui est considéré aujourd'hui comme une valeur repère pour les investigations projetées.

ESLO 2 a été conçu pour préfigurer la référence attendue dans un domaine qui en est encore à se structurer et dans lequel se manifeste de manière récurrente une demande de définition pour un format standardisé de *collecte*, de *conservation*, de *traitement* et d'*analyse* :

- la *collecte* sur le terrain est première, non seulement dans ses aspects techniques, aujourd'hui bien maîtrisés, mais dans la définition du profil de l'échantillon représentatif et dans la problématisation des interactions entre les témoins et les enquêteurs ;
- la *conservation*, qui inclut la préservation des supports, l'indexation des contenus et l'accessibilité (c'est-à-dire la protection) des données, conditionne le partage des sources à des fins d'étude scientifique ou didactique ;
- le *traitement*, en lien étroit avec le développement des matériels et des langages informatiques, suppose la maîtrise d'une chaîne d'opérations, depuis la conversion numérique des enregistrements jusqu'à une transcription balisée et ouverte à l'ensemble des interrogations pertinentes pour les demandes du linguiste, du sociologue ou des décideurs, des didacticiens voire du grand public ;
- l'*analyse* constitue l'épreuve des théories (et des logiciels) puisqu'elle compare les formalisations et les opérations et qu'elle valide ou infirme les hypothèses en prenant argument de leur compatibilité aux faits.

Les acquis en matière de conservation, de traitement et d'analyse seront reportés sur ESLO 1 comme le requiert la comparabilité attendue.

C'est le second objectif : évaluer la dynamique sociolinguistique du français en bénéficiant de la constitution d'un observatoire constitué par le renouvellement de l'enquête à quarante années de distance. Les usages de la langue ont évolué (vitesse d'élocution, contours prosodiques, tours syntaxique, choix des métaphores et des modalisateurs, stock lexical et modes de dérivation, termes d'adresse ou d'insulte, locutions et allusions...) : il reste à apprécier l'importance, la cohésion et l'orientation des phénomènes avant d'en proposer des principes d'explication. Le changement affecte aussi les jugements sur l'emploi : perception des « accents », tolérance au français familier et aux termes argotiques, relativisation des jugements scolaires. Autant qu'une pratique linguistique qui se transforme, il y a, dans l'usage de la parole, le modèle prégnant des régies de discours qui formatent de nouvelles normes et de nouvelles relations aux usages (modèles des animateurs télévisions et radios, des acteurs ou des professionnels de la parole publique, voix synthétiques des automates et boîtes vocales...) et qui tendent à établir les modèles de référence pour une communauté d'auditeurs en interaction.

La disponibilité des informations concernant les locuteurs d'ESLO 1 permet de prendre en compte la diversité des changements en les rapportant à leur époque et aux paramètres sociaux qui les déterminent, dessinant l'inégalité des résistances ou des propensions à la transformation de la langue, mais aussi une typologie et une dynamique de ses évolutions.

1.3 Lormont : variations sociolinguistique, dialectale et diachronique

Entre 1980 et 1981, une grande enquête sociolinguistique sur le français parlé a été menée à Lormont, communauté urbaine située au nord-est de l'agglomération de Bordeaux. L'une des variables analysées concernait l'élision ou le maintien du schwa en fin de polysyllabe, phénomène particulièrement pertinent dans le parler méridional.

Avant de procéder à une analyse sociolinguistique de faits de variations avérés comme celui du e muet, le recueil des données a été précédé d'une enquête du contexte social et pragmatique de Lormont au moyen d'une méthode statistique et ethnographique (constitution d'un échantillon de locuteurs représentatif de la structure socio-démographique de la communauté et recherche concernant la structure objective de l'espace social avec les représentations que se faisaient les agents sociaux de leur environnement)..

Cinquante-six personnes, natives de l'agglomération bordelaise, ont été retenues pour l'analyse sociolinguistique menée sous forme d'interviews de deux heures chacune, avec fiche biographique, questionnaire sociologique (suivant le modèle de Bourdieu) et linguistique.

Recueillir des données en reprenant le même échantillon de population et le même format d'interview que ceux de l'enquête de 1980 permettrait de mener une véritable enquête longitudinale, ce qui n'a jamais été fait dans l'histoire de l'étude du français parlé. Les avancées scientifiques d'une telle enquête sont potentiellement considérables : elle permettrait de mieux comprendre les processus d'évolution des variables étudiées et, par là-même, certains aspects du fonctionnement du langage comme les conditions d'émergence de patterns grammaticaux, pragmatiques, et phonologiques avérés.

Aujourd'hui, il importe d'élargir les variables observées grâce à de nouvelles modalités d'enquêtes centrées sur l'acquisition de données interactionnelles en intégrant des aspects interactionnels et conversationnels. Ainsi de certains phénomènes conversationnels comme l'autocorrection ou les pré-séquences, les écarts entre les styles formels et familiers. On obtiendra ainsi un deuxième corpus pour la comparaison dialectale avec une variation diachronique intégrée.

2. Les fonctions d'un corpus : standard, norme, benchmark

L'inscription d'ESLO 2 dans une démarche qui vise à définir les propriétés d'un corpus de référence dans la méthodologie de constitution de données implique la confrontation des différentes étapes de l'élaboration à celle d'autres programmes, aboutis ou en construction, afin d'anticiper les contraintes de l'interopérabilité tout en évaluant leur impact sur l'analyse et les théories linguistiques.

2.1 Vers la conception d'un prototype

La réalisation du corpus de référence est effectuée en trois temps :

- 1) une recension des pratiques en vigueur dans la communauté scientifique pour ce qui a trait aux corpus oraux (avec une participation au programme CNRS concerné, à l'ILF et à la coordination du *Guide des bonnes pratiques*),
- 2) une synthèse des recommandations élaborées par l'échange et le consensus à l'intérieur de la communauté des usagers et des experts exposés au projet, une démarche calquée sur les procédures de standardisation internationale.
- 3) une conception de traitement d'ESLO 1 et de fabrication d'ESLO 2 qui puisse permettre de les établir comme corpus prototypique.

Le premier principe retenu revient à considérer qu'un corpus numérique n'existe que par l'exploitation qui peut en être faite, aujourd'hui et dans l'avenir (ce qui suppose d'éviter toute opération irréversible qui n'autoriserait pas une récupération des sources dans leur état premier). Il faut déterminer au préalable et *a priori* l'ensemble non dégradant d'interrogations et de sollicitations, quels que soient les auteurs des demandes et leurs requêtes. Le conditionnement le plus aisément manipulable est celui obtenu par numérisation, c'est-à-dire par une transformation des data qui permet de tracer des graphes de relations. L'accessibilité, élargie au-delà des producteurs, requiert une paramétrisation qui autorise la consultation par des systèmes informatiques différents, sur support fixe ou à distance. La question de la structuration de la base pour l'adapter aux normes, standards et routines pratiqués est déterminante.

C'est dans cette phase qu'intervient la chaîne des actions qui vont de la capture du matériau sonore à la réponse à une demande qui peut être formulée par n'importe quel acteur de la recherche, qu'il interroge les propriétés physiques du signal ou l'emploi des connecteurs, le témoignage du consommateur ou les états mentaux du locuteur. La disparité des requêtes prescrit le choix de systèmes ouverts, évolutifs et interopérables, c'est-à-dire qui respectent les normes de transport et de lecture en vigueur et qui anticipent les normes futures. L'association avec d'autres équipes confrontées aux mêmes problèmes, et partant le choix des laboratoires, ont été dictés par cette perspective.

Un corpus numérique doit être associé aux protocoles et aux outils qui en permettent le stockage, le transfert et l'utilisation : que le rendu se présente sous forme de texte (transcription) ou de son (signal), il s'agit dans les deux cas de concevoir la structuration spécifique d'un objet. Les corpus numériques ont ceci de particulier que, par leur taille exponentielle et leurs possibilités rapides d'évolution ou d'obsolescence, ils ne se prêtent pas immédiatement à une exploitation spontanée et pérenne. Dans le cas d'un corpus sonore, les difficultés augmentent avec la spécificité des données à

manipuler. Comment articuler le système documentaire propre à l'exploitation du corpus et les objets dont il est constitué ? Les découpages ou les formats de requête, qui sont d'évidence sur du texte, ne le sont plus dès que le domaine d'investigation est un fichier sonore.

La finalisation de la recherche vers la définition d'une structure versatile et utilisable d'un corpus suppose au préalable une bonne représentation des objets à indexer, aussi bien en termes de typologie que de contenu. Un corpus numérique ne saurait se concevoir indépendamment des outils de traitement spécifique des métadonnées, en renfort des outils destinés au traitement même des objets. Avec un corpus de dix millions de mots, les opérations de structuration et de mise à disposition exigent que soit établi en amont un système d'indexation des données au moyen d'un étiquetage par métadonnées. Ce sont les normes et les standards qui se mettent en place dans ce domaine, autour du XML, langage de structuration des données, et de ses dérivés.

2.2 Les corpus linguistiques

Dans les disciplines linguistiques, les corpus oraux représentent à la fois une méthode de travail et un outil d'étude. Au fur et à mesure des développements informatiques, les techniques de numérisation et les possibilités de diffusion par le Web rendent possible une gestion raisonnée et à faible coût des corpus. Le nombre croissant de laboratoires qui ont fait le choix de pratiques orientées vers une exploitation automatique des ressources et la nécessité de transférer les données stockées sur support analogique a décidé la mise en place de vastes programmes de numérisation.

Les corpus sont devenus la norme de fabrication et de constitution des équipements de la langue (système phonologique, grammaires et dictionnaires, répertoires de locutions, analyseurs morphologiques...) en même temps que les premiers auxiliaires de la recherche en sciences du langage. Il s'en est suivie une multiplication, à l'échelle internationale, des programmes consacrés à la confection de corpus représentatifs, avec toutes les interrogations qui découlent d'une réflexion sur la qualité de la correspondance entre l'échantillon et la masse de données que serait l'impossible projet d'une collation exhaustive des échanges langagiers. Ainsi, la multiplication des entreprises de fabrication de corpus oraux pour la seule discipline que constitue la linguistique a en quelques années abouti à une importante disparité des codages, des formats et des méthodes de catalogage, de référencement et de diffusion.

Cette variété de procédures aboutit à un gaspillage d'investissement, matériel et humain, dans la mesure où l'effort de codage, de formatage et de développement est accru d'autant par la quantité d'initiatives, un coût encore augmenté par le nombre de passerelles à définir afin de rendre interopérables différentes initiatives incompatibles dans leur philosophie et leurs technologies alors même que certaines applications nécessitent la compilation de ressources provenant de différentes sources pour obtenir les volumes critiques nécessaires. À plus long terme, le cumul des connaissances comme le transfert vers d'autres domaines, aussi bien académiques qu'industriels, est rendu aléatoire.

Les descriptions et les besoins sont en effet différents pour des sociolinguistes, des linguistes de terrain, des dialectologues, des phonéticiens, des spécialistes de TAL ou d'ingénierie linguistique, sans parler d'exploitation en sciences humaines et sociales voire en sciences formalisées. Les analyses linguistiques et donc les codages varieront selon les investigations, allant d'études dédiées à des phénomènes phonologiques particuliers jusqu'à des exploitations qui se satisfont de transcriptions orthographiques aménagées, ou qui au contraire incluent des transcriptions phonétiques, des découpages morpho-syntaxiques, des lemmatisations, des alignements texte/oral, etc.

3. Phasage

On distinguera, dans le programme soumis à l'ANR, deux types de corpus : d'une part, ceux qui sont déjà constitués et qu'il s'agit de reprendre pour les réinscrire dans les standards d'aujourd'hui et les insérer dans les circuits d'échange télématiques. D'autre part, un corpus spécifique à élaborer à partir de principes qu'il convient de tester afin d'en vérifier la pertinence et la faisabilité.

Concernant les corpus constitués, on mentionnera les fonds disponibles auprès des laboratoires associés au projet : outre ESLO 1 (voir infra), le corpus de Lormont et ceux qui pourront être mis à

disposition par Modyco (Université Paris X), ceux du LACITO et du CORAL (créoles à base portugaise, palikur, wayana et wayampi en Guyane).

Le projet central autour duquel s'articulent les différents modules est néanmoins la réalisation d'un corpus spécifique, ESLO 2, dans sa visée cumulative (ajouter des données nouvelles à d'autres qui leur sont comparables) et réflexive (accompagner l'enquête, le traitement et l'exploitation d'une analyse de l'expérience pour participer à la définition des normes).

Bien que chaque étape suppose une anticipation sur le déroulement du processus dans son entier et que toute réalisation ne puisse être conçue sans un effet en retour sur les travaux réalisés dans les phases antérieures, on distinguera à la suite les différents temps de fabrication d'ESLO 2 en rappelant que l'objectif premier est la mise au point d'un corpus oral qui puisse être rapproché d'ESLO 1 tout en servant de prototype aux initiatives qui sollicitent des données orales pour l'élaboration de connaissances.

Phase 1 : Exhaustivité, représentativité, proportionnalité

A la différence de l'écrit public (imprimé) qui est centralisé par de grandes institutions en sorte que la conservation est en théorie identique à la production, aucune instance ne peut préserver l'intégralité des productions orales. De plus, alors que l'écrit privé suppose un fort degré de contrôle et une certaine mise en situation liée à la décontextualisation de la production et de la réception, l'oral se situe en dehors des processus d'appropriation contrôlé, même si l'école exerce une influence sur son emploi, et les conventions d'emploi n'y ont pas la même incidence. L'intérêt pour les écritures ordinaires, les chats ou le SMS a pu fragiliser des distinctions établies, elle ne les a pas abolies.

A défaut d'une exhaustivité des productions, c'est une exhaustivité des usages qui doit être envisagée, dans les différentes dimensions où ils s'inscrivent, c'est-à-dire au cœur même de la première difficulté que soulèvent ces corpus : la variation. ESLO 1 avait exclu deux catégories de locuteurs : les enfants (0 – 20 ans) et les non francophones. Si le contact de langues est un cas très particulier qui suppose une enquête spécifique, la nécessité d'un groupe témoin enfant/adolescent est indispensable.

De façon générale, poser la question de la représentativité, c'est d'abord procéder à l'inventaire des causes de variation. On les distribuera en trois catégories :

- celles qui sont liées à des propriétés intrinsèques des locuteurs, avec une ventilation par âge, par sexe, par CSP, par trajectoire sociale et par origine géographique, sur le modèle de ce qui avait été fait avec ESLO 1 et qui bénéficie depuis des apports cumulés de l'anthropologie et de la sociologie (on intégrera, en face du CSP, la définition donnée par le témoin lui-même de son statut social) ;

- celles qui relèvent des situations de discours et des niveaux de langue exploités, y compris les actes performatifs et la construction des identités relationnelles, la définition du cadre, l'organisation des tours de parole, la construction du point de vue, saisis dans la variété des situations de collecte et des modes de relation entre enquêteur et enquêté ;

- celles qui articulent formats cognitifs (type de tâche à accomplir : récit, description, suggestion, plaisanterie, demande d'assentiment) et exploitation des ressources linguistiques (figures du discours, opérations syntaxiques, choix lexical, deixis et procédures anaphoriques...).

Ces trois versants, distingués pour la commodité de l'exploitation, sont corrélés. Ils se caractérisent par leur caractère obligé (tout discours en tant qu'il est oralisé est assignable à un locuteur, un contexte de production et une tâche cognitivo-linguistique). A cette étape, il s'agit de préfigurer une grille de propriétés, comprenant un nombre fini d'entrées, ouvrant à une normalisation des types de requêtes possibles. Si un siècle et demi d'enquêtes sociologiques ont abouti à un ensemble réglé de conventions concernant le premier point, il n'en va pas de même pour les deux suivants qui ont fait l'objet de nombreuses études sans converger vers un catalogue fini d'items. La démarche retenue consiste à partir de cette interrogation : quelle question peut être formulée qui éclairerait les données de les appréhender à partir de telle caractérisation ?

ESLO 2 a fait appel à une sociologue, M. de Saint-Martin (EHESS), pour évaluer ses propositions de questionnaire et les échanges ont mis en évidence que la diffraction des modes d'opérabilité de la langue en situation ne se prêtent pas si facilement à une homogénéisation ou à une codification. A partir d'un éventail de situations et d'emplois, de façon prédictive et rétroactive, il est

prévu de procéder à une extension catégorielle permettant la détermination générique d'un ensemble fini de propriétés exploitables pour le catalogage.

La question d'une représentativité quantitative (présence dans le corpus d'une proportion homologue de types d'emplois et de locuteurs, conforme à celle des usages quotidiens et des rapports entre groupes sociaux) est pour le moment indécidable, sinon par la négative : on peut être assuré qu'il existe moins de déclamation poétique que de demandes de renseignement dans les magasins d'informatique. Une esquisse de ces mesures est néanmoins envisagée dans la sollicitation d'une reconstruction des successions d'interaction quotidiennes.

Phase 2 : Techniques de collecte : formats d'enregistrement et numérisation

Cette phase est aujourd'hui la mieux maîtrisée encore qu'elle soit susceptible d'évolutions extrêmement rapides. Les équipements proposés par les fabricants - qui cherchent souvent à imposer leurs propres normes dans une logique commerciale d'exclusivité - permettent un échantillonnage du signal, un stockage et des formes de restitution dont la qualité dépend plus souvent des conditions d'enregistrement que des performances électroniques. Les moyens de capture (micros) et de conversion ne sont pas au centre de nos interrogations même s'ils feront l'objet d'un examen attentif, en particulier avec les institutions (BnF) et les associations (AFAS) impliquées dans une conception académique de l'acquisition de données sonores.

A cette étape, il s'agit plus d'une sollicitation de compétences pour mesurer les avantages et les difficultés inhérents à chaque solution technique que d'une véritable normalisation. On trouve dans les fiches techniques du *Guide des bonnes pratiques* des indications sur ces points, en particulier sur l'obsolescence des formats et la durabilité attendue des produits, leur capacité de transfert vers d'autres supports et la fiabilité de leur stabilisation (qualité du signal) dans le temps.

Phase 3 : Formation des enquêteurs et information des témoins

Alors que les pratiques d'enquête sont aujourd'hui de règle dans tous les secteurs des sciences humaines et sociales, il n'existe pas véritablement de définition satisfaisante concernant ce que devraient être les critères d'une enquête valide. La formation se fait sur le tas et les praticiens apprennent autant et plus de leurs échecs (appareil d'enregistrement défectueux, questions oubliées, destruction des données au cours des manipulations, mauvais étiquetage...) que des conseils qui leur ont été prodigués.

ESLO 1 a réuni des collaborateurs de bonne volonté. Les enquêtes du CELITH et du LACITO ont été réalisées par des semi-professionnels (étudiants en thèse) ou des linguistes de terrain. Au CORAL, certaines études ont été conduites par des chercheurs immergés in situ sur de longues périodes (Guyane, Guinée Bissau et Cap-Vert, France), d'autres par des missions de durée variable.

Il paraît essentiel d'intégrer dans les critères de variation celle qui se trouve liée à l'enquêteur, à ses propriétés d'âge et de niveau social, de familiarité avec le témoin et de connaissance des objectifs de l'enquête. Dans un module d'ESLO 2, consacré au réenregistrement d'une dizaine de témoins d'ESLO 1 quarante années plus tard, l'enquêtrice est elle-même une retraitée de l'enseignement qui, pour des raisons d'âge, est plus à même de dialoguer avec des personnes âgées de plus de soixante ans.

On distinguera quatre types d'enquêteur :

- enquêteur « professionnel », rompu aux techniques de l'interview et de la collecte des documents, généralement par une pratique éprouvée au fil des ans ;
- enquêteur « apprenant », correspondant à la situation d'un étudiant en cycle Master ou doctorat ;
- enquêteur « médiateur », sollicité pour aller recueillir des données auprès de locuteurs auxquels l'enquêteur ne peut avoir un accès direct pour des raisons de langue ou d'organisation sociale (public féminin quand l'enquêteur est masculin par exemple) ;
- enquêteur « involontaire » lorsqu'une conversation est intégrée à un corpus sans que la finalité première ait été conçue au moment de la confection de l'enregistrement (lorsqu'un étudiant apporte une cassette où il parle avec quelqu'un de sa famille par exemple).

S'il est difficile d'imaginer un protocole qui permettrait de reproduire exactement la même forme d'entretien en variant seulement l'enquêteur, il est concevable de rechercher des indices, dans les formes d'adresse, le registre lexical, les marques de connivence ou de distance, qui répercuteraient la diversité des relations d'enquête.

L'objectif de cette phase serait, en partenariat avec les équipes impliquées sur différents terrains, et si possible à l'échelle européenne, de prévoir une école d'été qui exposerait les différentes pratiques, avec des exemples précis de résultats, qui réunirait les informations documentaires, en particulier électroniques et qui assurerait en parallèle une formation de haut niveau dans l'enquête. C'est une perspective ambitieuse et qui suppose l'appui des institutions (CNRS, Max-Planck Institut, ESF...) : elle a été anticipée par des séminaires et des sessions de travail qui acquerraient une autre dimension.

Phase 4 : Recueil des données

On ne traitera pas dans cette partie des données déjà collectées qu'il s'agit de récupérer pour les inscrire dans des formats exploitables par la communauté. L'enquête constituée en prototype est ESLO 2. Il s'agit, en deux années maximum, à partir d'un questionnaire réfléchi et d'un échantillon de locuteurs et de situations (phase 1), avec un matériel fiable d'enregistrement et de stockage (phase 2) et des enquêteurs préparés à leur tâche (phase 3), de recueillir 400 heures de parole, quarante pour cent par entretiens semi-directifs, quarante pour cent par collecte dans des situations variées, le restant par l'adjonction de modules spécifiques. Au nombre de ceux-ci, on citera :

- un module « témoins ESLO 1 » : cf. supra.
- un module « enfants » qui n'a pas d'équivalent dans ESLO 1 et qui pourrait faire l'objet d'une saisie vidéo,
- un module « personnalités » constitué par des interviews de personnalités de la société civile, de politiques et de publicistes qui donneront une image sonore de la parole officielle, commune à tous les auditeurs (ce sont ceux qu'on entend dans les médias, les cérémonies officielles, etc.).

A chaque élément d'enquête sera associé une méta-donnée concernant les circonstances de l'enquête (mode de contact du témoin, négociation de l'entretien, caractéristiques sociales de l'enquêté et de l'enquêteur, format d'échange, incidents...) qui feront l'objet d'une saisie séparée avec une possibilité d'interrogation distincte.

La collection des données et des méta-données constituera le matériau d'ESLO 2 qui, pour devenir le corpus de référence attendu, devra faire l'objet de la série d'opérations énumérées dans les phases 5 à 12, parallèlement à des traitements similaires effectués sur ESLO 1 afin d'assurer la comparabilité des data et, autant qu'il sera possible, sur les autres corpus apportés à ce programme de l'ANR en provenance d'Afrique, d'Amérique et de France.

Phase 5 : Codage et catalogage

Ces deux opérations sont liées aux propositions formulées dans le cadre de CatCod, créé en 2005 par une communauté d'enseignants-chercheurs, de chercheurs et d'ingénieurs travaillant sur des corpus de parole afin de préfigurer des solutions efficaces de partage des données, c'est-à-dire la définition de standards et de formats d'échange. La première initiative a consisté dans un échange d'informations sur les différentes pratiques en usage, structuré sous la forme d'une structure d'Equipe Projet Multi-Laboratoire du CNRS, l'EPML-50, « Corpus d'interaction langagière ». Des outils collaboratifs ont été conçus afin d'organiser et d'animer la communauté autour de ces thèmes. Il y a à présent une demande pour que le travail ne soit plus seulement descriptif et rétroactif mais prescriptif et prospectif. En ce sens, l'acquis de l'EPML-50 est directement associé au traitement d'ESLO 2 et au recyclage des autres corpus associés à cette demande d'ANR. Terrain d'expérimentation idéal, parce qu'il associe dès le début de sa conception les acteurs de l'EPML-50, ESLO 2 est déclarée comme membre participant du nouveau centre de ressources CRDO au sein d'un réseau de besoins applicatifs.

Cette phase doit aboutir, en partenariat avec les équipes impliquées au CRDO et à l'EPML, à la rédaction d'un document de spécifications qui propose des normes de standardisation, avec une large

consultation et une recherche du consensus entre les participants. Elles auront été utilisées et éprouvées, à titre expérimental, dans la reprise d'ESLO 1 et la confection d'ESLO 2 en matière de codage et de catalogage.

Afin de stabiliser les acquis de cette recherche, le document de spécifications sera transmis au consortium « Text Encoding Initiative » (TEI) à titre de proposition pour le chapitre transcription de l'oral. La prise en compte de nos travaux dans le cadre normalisateur de la TEI apporterait une garantie de pérennité et l'assurance d'une diffusion, participatif à la réduction d'une diversité des usages préjudiciable à l'échange et à la conservation.

Phase 6 : Transcription et alignement

Dans la continuité du travail de codage, la transcription pose une question spécifique pour laquelle des réponses différentes ont été apportées. En partant des acquis de tous les laboratoires, y compris ceux du LACITO, du CELTH, du LI et du CORAL, il sera fait, de la transcription des quatre cents heures de parole, un benchmark des solutions qui sera indéfiniment réutilisable pour d'autres applications.

Un problème particulier concerne l'alignement du signal sur la transcription, dans le prolongement des études entreprises à Leuven. Le format d'analyse du signal requiert une réflexion sur les outils disponibles, les manipulations effectuées, les distorsions qui peuvent en résulter et l'exploitabilité des représentations et des affichages. La possibilité de retrouver, à partir d'un item ou d'une séquence, la représentation temporelle ou spectrographique, le contour mélodique ou les variations de l'énergie est possible mais sur des formats qui exigent de grosses capacités de mémoire puisque interfèrent une recherche dans les données, un alignement de deux fichiers et du traitement graphique. La mise au point de systèmes à différents degrés de granularité, et la possibilité d'une sélection des informations, devrait permettre d'accélérer les temps de traitement et d'affichage en éliminant ce qui, pour une interrogation spécifique, n'est plus de l'information mais du bruit.

En revanche, l'opération inverse (retrouver l'ensemble des séquences présentant un certain type de variation en F0, ou tous les moments réunissant les formants d'une voyelle donnée) reste encore très complexe puisque la requête ne porte plus sur une chaîne de caractères mais sur la numérisation de données physiques. C'est un aspect qui ne sera pas traité dans cette étude.

Phase 7 : Etiquetage, catégorisation et lemmatisation

L'étiquetage n'est pas seulement un problème technique. S'il existe des logiciels performants, avec des taux de reconnaissance assez satisfaisants, leur entraînement a surtout été éprouvé sur des corpus écrits. Une caractéristique du français parlé est la présence d'unités non standard (apocopes et aphérèses, diminutifs, expression argotiques, troncations, « fautes de français », duplications, néologie dérivationnelle, antonomase...) et les attestations de dérivation impropre et de polycatégorialité qui brouillent les caractérisations taxinomiques.

A côté d'un travail technique de balisage des corpus et de lemmatisation, c'est donc une réflexion sur la morpho-syntaxe du français qui est à conduire. Une autre application concerne la détection des entités nommées dont la définition est à étendre à celle d'une détection des entités dénommantes.

Phase 8 : Anonymisation

L'un des obstacles récurrents qui se présente dans l'exploitation collective des corpus oraux concerne la protection des personnes qui ont accepté d'être sollicités et de répondre. Comme le prévoient les engagements pris par le chercheur, sous forme contractuelle, l'enquêté a le droit de conserver son anonymat ce qui suppose l'effacement ou le bipage de toutes les indications qui permettraient de lever l'anonymat. Il en va de même pour les personnes mentionnées au cours de l'interview qui sont en droit d'exiger la protection de leur vie privée.

Dans ce cadre, et en prolongement des recherches entreprises dans la base 7, ESLO 1 et ESLO 2 doivent servir de test aux programmes développés dans le cadre du LI concernant la reconnaissance des

entités nommées. En recourant à CasSys, un outil qui traite des corpus Unitex et qui permet, entre autres, de détecter les noms propres avec un taux d'efficacité supérieur à 90 %, on soumettra les résultats obtenus aux corpus oraux avec deux objectifs :

- tester sur du français oral transcrit des outils élaborés sur du français écrit,
- accroître la définition des entités nommées à des catégories indépendantes du nom propre (le maire d'Orléans en 2000 est l'exact équivalent d'un nom propre d'un point de vue juridique et, peut-être, philosophique).

Sur les centaines d'heures à traiter, on peut concevoir que tous les problèmes de l'anonymisation seront représentés ce qui permettra la confection d'un vade mecum des éléments à prendre en considération à cette étape du travail :

- détection et sélection des noms propres (tous ne sont pas à anonymiser : la Loire et Jeanne d'Arc ne sont pas à inclure dans l'effacement),
- repérage des éléments d'identification hors nom propre (profession, lien de parenté, participation à un fait divers),
- coordination à distance des éléments d'identification (un professeur ne permet pas d'identification, il n'en va pas de même s'il est précisé par ailleurs que c'est un professeur d'Université, spécialisé en électronique et que c'est une femme, auquel cas on peut arriver à un singleton).

Phase 9 : Stockage, archivage et indexation

La conservation des données, comme leur saisie, a fait l'objet d'un travail de réflexion dont on trouve la synthèse dans le *Guide des bonnes pratiques* dont les prescriptions seront suivies. Les questions du choix de l'institution d'archivage (le laboratoire lui-même, un centre spécialisé à créer sous l'égide du CNRS, les Archives départementales ou nationales, la BnF, un serveur dévolu à cette tâche) et des autorisations et des modes d'accès seront tranchées à la fin de l'ANR avec une justification de la solution retenue. Les discussions en cours à l'intérieur de chacun des établissements ne permettent pas d'anticiper une décision même si l'engagement des chercheurs dans les initiatives CNRS leur fait apprécier positivement une solution d'attente qui préfigure une réflexion en chantier à l'intérieur du Département des archives sonores de la BnF.

La question de l'indexation est autrement plus complexe. En effet, à la différence d'un ouvrage que l'on peut feuilleter, d'un fichier texte que l'on peut consulter par appel d'items, l'indexation des fichiers sons pose un problème dès lors qu'on ne se contente pas d'en traiter la partie transcription mais qu'on se propose d'en relever les caractéristiques de signal. Une proposition sera faite concernant la production de différents niveaux d'indexation :

- un code alphanumérique d'identification,
- une fiche de métadonnées (voir phase 4),
- un descriptif court du contenu (une centaine de signes),
- une synthèse d'un millier de signes recensant (i) les caractéristiques de l'enregistrement, (ii) les principaux éléments de contenus avec une liste de mots-clés réunis dans un lexique séparé, (iii) les problèmes que présente l'enregistrement,
- une fiche évolutive recensant la série des interventions opérées sur le fichier et l'historique des requêtes,
- un dossier associé concernant les exploitations du document dans la littérature scientifique et les applications.

Phase 10 : Mise à disposition : construction du site et traduction

Pour que ESLO 1 et ESLO 2 puissent effectivement fonctionner comme références dans la constitution des corpus oraux, en France et à l'étranger, il faut que soient rendues simultanément disponibles les données et l'explicitation des opérations qui les ont configurées dans les formats auxquels elles seront accessibles.

L'accès suppose l'installation (et la maintenance) d'un site librement consultable par Internet, avec une convivialité et une ergonomie des applications. Il ne s'agit pas seulement d'empiler des heures de transcription mais de réfléchir à la façon dont les interrogations, et pas seulement celles des linguistes, seront adressées. Le principe de construction du site sera fondé sur la notion d'usage, en conciliant, dans une réflexion sur l'interface graphique et les parcours possibles, l'hétérogénéité des sollicitations :

- celle des chercheurs en science du langage à la recherche d'exemples, d'occurrences ou de statistiques (quelle proportion de propositions négatives, quel pourcentage de métaphores, quelle corrélation entre telle catégorie de noms et telle catégorie d'adjectifs (« tous les noms d'animaux » ou « toutes les couleurs » suppose la possibilité de classements hyperotaxiques), combien de hiatus par tranche de mille voyelles réalisées.. ?),
- celle des chercheurs en sciences humaines qui, pour des interrogations sur l'interaction, les groupes sociaux, les types de communication ou les représentations mentales auront recours soit à des indices lexicaux, soit à une requête sur les méta-données,
- celle des étudiants qui pourront trouver, grâce à la dimension réflexive du travail, des moyens d'autoformation,
- celle des décideurs, en particulier en matière de politique linguistique,
- celle des développeurs dans les industries de la langue et les applications informatiques et didactiques, de FLE comme de FLM.

Ce site, s'il est prioritairement destiné à la communauté des chercheurs de langue française, est aussi voué à servir de vitrine internationale. La question d'une traduction des présentations et des opérations est d'ores et déjà posée, d'abord en anglais mais aussi en allemand et en espagnol pour intégrer les compétences et les collaborations déjà présentes au sein des équipes. Il ne s'agit pas seulement de traduire des mots mais d'articuler à l'expérience et aux acquis de chaque communauté scientifique les résultats propres produits par ESLO dans les paramètres particuliers de la situation du français en France.

Phase 11 : Données partagées : interopérabilité et protections

Parallèlement à ce que peut assurer une mise à disposition la plus ouverte possible (avec les restrictions qui s'imposent concernant la protection des personnes et des fonds), ESLO 1 et ESLO 2 doivent pouvoir être inter-opérables avec les corpus linguistiques élaborés par les chercheurs. Une première étape de l'interopérabilité sera réalisée dans la confrontation avec des corpus présents dans les laboratoires participants : corpus palikur, corpus de créoles, corpus de français d'autres régions. Ces rapprochements auront valeur de test pour une extension à l'ensemble des données sonores recueillies à des fins d'analyse linguistique. Ils représentent d'ores et déjà cinq états de différenciation :

1. corpus de français issus de la même région et recueillis dans les mêmes conditions à quarante années de distance (ESLO 1 et 2),
2. corpus de français issu d'une autre région et recueilli avec des méthodes et des perspectives variationnistes différentes (Corpus Lormont),
3. corpus de français écrit (journaux),
4. corpus de langue orale à base romane avec variation interne forte (créoles de Guinée Bissau, du Cap-Vert et de Sao Tome et Principe),
5. corpus de langue orale sans apparentement avec le français sinon par contact (palikur, wayana, wayampi).

Les questions qui se poseront dans l'interopérabilité de ces corpus auront valeur d'exemple pour les autres corpus qui pourront y être associés ultérieurement.

Phase 12 : Applications et développements

Les études sur corpus concernent directement le Traitement Automatique des Langues Naturelles (TALN). Comme l'ont montré par exemple (sur le français) les travaux du LIUM (Le Mans), du DELIC (Aix-en-Provence), du CLIPS-IMAG (Grenoble), du LORIA (Nancy) ou du LI (Tours-

Blois), la recherche d'une description linguistique de l'oral sur des corpus pilotes permet d'orienter les recherches et le prototypage des systèmes de traitement automatique du langage parlé (compréhension de parole, dialogue oral homme-machine, traduction parole-parole etc.). C'est cette perspective, à l'interface entre sciences du langage et TALN, que nous comptons en profitant des caractéristiques des corpus ESLO.

Dans le cadre de ce projet, nous allons plus précisément nous intéresser à la description de deux phénomènes langagiers en situation d'élocution oral spontanée :

- disfluences de l'oral, hésitations, répétitions et auto-corrections,
- co-référence anaphorique en situation de parole spontanée.

Le traitement de ces procédés est essentiel dans le développement d'applications orales interactives.

Les disfluences de l'oral spontané sont susceptibles de perturber fortement la structure syntaxique des énoncés oraux, ce qui rend difficile leur analyse automatique. Il est donc important de les décrire linguistiquement dans l'objectif d'atteindre une modélisation suffisamment robuste pour être implémentée sur des systèmes opérationnels. Des recherches linguistiques sont menées dans cette optique depuis maintenant une vingtaine d'années. En terme d'application, celles-ci se réduisent cependant le plus souvent au dialogue oral finalisé sur des tâches bien précises (réservation aérienne par exemple) et la communauté scientifique ne dispose pas de connaissances précises sur le degré de généralité de ces résultats. Il semble que le contexte d'interaction (degré d'interactivité, familiarité des interlocuteurs, degré de spécialisation du dialogue) influe d'une manière sensible sur la distribution des procédés utilisés. Les corpus ESLO offrent des garanties en terme de significativité des données et surtout de contrôle des variables pouvant influencer sur ces procédés. A partir de ces nouvelles ressources, on se propose de répondre à ces deux questions : (1) quelles sont les variables sociolinguistiques, notamment les différentes caractéristiques du contexte d'interaction, qui influent sur le taux et le mode de réalisation des disfluences orales, (2) les résultats obtenus sur de la parole conversationnelle peuvent-ils être étendus à d'autres genres / registres de langue, au sens de Biber

La coréférence anaphorique est un autre problème qui touche aussi bien à l'écrit qu'à l'oral et le calcul de la référence est ainsi un problème de première importance pour le dialogue oral homme-machine. Dans le cas d'un serveur vocal d'information, il est par exemple essentiel d'accéder à l'ensemble des objets de l'application auxquels réfère l'utilisateur dans sa requête avant de pouvoir interroger la base de données. Consacrées avant tout à l'anaphore pronominale, les recherches actuelles sont limitées par leur application restreinte à l'écrit ou au contraire à un dialogue oral très finalisé. L'étude des anaphores pronominales sur le corpus ESLO traitera les questions suivantes : (1) dans quelle mesure les heuristiques de résolution développées pour l'écrit sont-elles applicables en oral spontané ? (2) Dans quelle mesure la théorie du centrage est-elle vérifiée sur l'oral conversationnel ? (3) La pertinence de cette théorie est-elle variable suivant le genre oral considéré ?

Les réponses à ces interrogations seront envisagées dans une perspective quantitative faisant appel à une analyse différentielle / variationniste entre différentes sous-collections du corpus ESLO.

Phase 13 : Mise en place du suivi (maintenance, jouvence et sécurité) et applications

La pérennité du corpus est liée à la capacité des participants de l'équipe à concevoir une maintenance du site, à une intégration progressive de nouveaux enrichissements des données et à l'adaptation aux nouvelles normes, notamment pour ce qui a trait à la sécurité. Dans le calendrier prévu, les applications et les développements, même s'ils font l'objet d'une réflexion en amont, ne pourront guère avoir dépassé le stade de projets. Ils sont néanmoins mentionnés pour ce qu'ils contribuent à leur façon à la définition du corpus, dans les sollicitations qu'ils représentent et dans les attractions qu'ils exercent. On pense en particulier aux applications dans le domaine du TAL (reconnaissance de la parole, interface homme / machine, analyseurs syntaxiques, constitution automatique de dictionnaires, traduction...) et à celles dans le champ de la didactique, les chercheurs du CORAL participant à des formations de Master et de doctorat associées à ces problématiques.

L'ensemble de ce travail s'inspire des apports de la linguistique qu'il doit à son tour enrichir. Le champ d'application va de la phonologie au TAL, en passant par la morpho-syntaxe, la lexicologie, la

sémantique et pragmatique, l'analyse de la conversation et du discours. Produit d'une formation discursive, le discours y sera envisagé comme ensemble de régularités morphosyntaxiques et verbales fonctionnant comme autant d'indices d'une distribution en différents sous-genres. Les formats d'échange et d'argumentation, les expressions figées et les emplois allusifs (polyphoniques) seront également intégrés dans une contribution orientée corpus à la théorie du langage.

Conclusion

Ce projet manifeste une ambition forte. Il était en discussion entre les partenaires depuis que ESLO 2 a été relancé et l'ANR peut apporter des moyens pour sa réalisation. Constituer un corpus prototype à toutes les étapes est une tâche ambitieuse qui prend sa véritable dimension dans le défi qu'elle relève : établir le plus grand corpus oral francophone diffusé, comparable, à l'échelle du français, aux états premiers du BNC et appelé à devenir, comme celui-ci, un élément du patrimoine historique.

Il s'agit d'apporter, à une base de corpus oraux, des matériaux dont la collecte et le traitement pourront soutenir la comparaison avec les exigences apportées par l'INaLF aux corpus écrits. Dans un délai de trois ans, et sous réserve de disponibilité des financements, VARILING doit (1) accumuler des heures de parole et développer une expérience critique sur leur mode de collecte et d'analyse, (2) créer des matériaux exploitables par les chercheurs et structurer la communauté « corpus », (3) ouvrir des champs d'investigation à l'ensemble des sciences humaines et apporter les attestations requises par la lexicographie ou l'élaboration de grammaires de référence. Le projet est aussi destiné à inaugurer une nouvelle politique de conservation du document scientifique sonore non musical et à préparer des collaborations dans les applications, didactiques et informatiques.

B-3 – Bibliographie et état de la question

La description des langues, dans la tradition linguistique, a été effectuée à partir de la langue écrite, normalisée, avec une valorisation des textes littéraires. Le premier changement apporté par les technologies informatique a été la constitution de grands corpus, le plus souvent étayés par de l'imprimé (cf. en France l'importance accordée au journal *Le Monde* dans les analyses et les corpus d'exemples). Lorsqu'il a paru nécessaire d'étudier la langue parlée, ce sont généralement des corpus de radio et de télévision qui ont été exploités, en dépit des contraintes de format et de sélection des locuteurs inhérentes à ces médias.

La constitution de grands corpus de français parlé non médiatisé s'est heurtée à quatre difficultés :

1. un problème de représentativité (en lien avec une sociologie des locuteurs et des situations de communication),
2. des interrogations concernant la collecte (difficultés d'enregistrement en milieu bruyé, statut juridique des documents, protection de la vie privée...),
3. des solutions à inventer pour le traitement de la variation qui, d'un paramètre, devenait un obstacle à la représentation, au classement et à l'exploitation des données,
4. des incertitudes concernant la recherche sur les corpus obtenus.

Certains pays ont pris une avance en ce domaine et si l'exemple le plus souvent cité est celui du BNC (British National Corpus), il existe d'autres travaux d'une qualité équivalente en Allemagne, Espagne (Catalogne), Pays-Bas... En français, il y a :

- soit des corpus partiels (Aix-en-Provence, Lyon...),
- soit des corpus thématiques (PFC),
- soit des corpus variationnistes extra-métropolitains (Québec).

Il s'agit de conjointre des expériences et des approches qui existent ailleurs, ou autrement ou encore sur d'autres langues à la situation du français, en partant d'une collecte dans une région peu dialectalisée et moins soumise à la pression internationale des langues que la région parisienne.

Le travail sur l'analyse de corpus en français parlé dans une triple dimension de variation (sociale, diachronique et dialectale) interagit avec l'élaboration d'outils et d'instruments du TAL. Il ne semble pas qu'un programme comparable existe, en France, que ce soit à des fins d'expertise ou d'application. La perspective qui se dessine est une coordination de la recherche à l'échelle des corpus de français en Europe (variantes régionales, Belgique, Suisse...) et dans le monde (en particulier, comparaison avec le Québec) et aussi une confrontation et un cumul des expériences et des entreprises engagées en Europe (avec une continuation possible sur des crédits européens).

ESLO

Formulaire témoin

Les informations nominatives ci-dessous ne sont pas publiques, elles seront stockées dans une base à part.

Nom :

Prénom :

Nom de jeune fille :
.....

Adresse :

Année de naissance :
.....

Sexe : M – F **Lieu de naissance** (ville + région/département) :
.....

NIVEAU D'ETUDES

- Primaire
- Certificat d'études
- Secondaire (collège, middle school)
 - Complet Incomplet
- Lycée
 - Complet Incomplet
- Supérieur
 - +1 +2 +3 +4 +5+

Commentaires
.....
.....

PROFESSION

1- Statuts :

- Professionnel actif
- Retraité
- Chômeur
- Autre

2- Catégorie professionnelle (Code INSEE 1) :

- agriculteur exploitant
- artisan, commerçant ou chef d'entreprise
- cadre ou profession intellectuelle supérieure
- profession intermédiaire
- employé
- ouvrier

- retraité
- autre personne sans activité professionnelle
-

3- Spécifiez la profession en vos propres termes :

.....

.....

LANGUE(S)

Langue(s) parlée(s) [Langue(s) parlée(s) (passive(s) ou active(s)) autre(s) que le français] :

.....

.....

Commentaire(s), niveau de langue(s) :

.....

.....

SITUATION FAMILIALE

- Marié
- Célibataire
- Concubinage (marital, pacs, etc.)
- Autre (veuf(ve), divorcé(e), en couple sans vivre sous le même toit).

Précisez.....

.....

Année d'arrivée à Orléans (si depuis la naissance, indiquez l'année de naissance) :

.....

Domicile actuel (Ville + quartier) :

.....

ENFANT(S)

Nombre :

Sexe, année de naissance, scolarisation et autres commentaires :

.....

.....

REMARQUE(S) DIVERSE(S) :

.....

.....

.....

.....

.....

.....

.....

LCO / PANGLOSS

Programme d'Analyse Géographique et Linguistique à Orléans : Sciences et Société

Objectifs et finalités

L'objectif est de construire, sous forme d'une monographie associant géographes et linguistes, une enquête au centième de la grande enquête INSEE / INED sur les langues parlées en France. La taille de l'échantillon et l'articulation au travail d'un laboratoire de linguistique permettront d'assurer une collecte et surtout un traitement qui affineront et approfondiront les résultats globaux déjà obtenus à l'échelle nationale. Le travail, qui comprendra une phase d'enquête et de transcription, une phase de constitution de bases de données et de cartographie et une phase d'analyses et de prestations, a pour objectifs, au-delà d'une connaissance réflexive des flux de population,

- de recenser une diversité patrimoniale,
- de raisonner l'apport culturel et linguistique des populations allophones,
- de mesurer la dynamique sociale des langues (transmission, intermariage, déplacements intra- et internationaux...)
- de reconstruire les réseaux d'échange et de production pour lesquels la langue constitue un facteur explicatif.

Programme de travail

- Préparation de l'enquête (constitution de questionnaire, canevas d'entretien, etc.)
- Enquête et identification des langues (recueil de textes en différentes langues et transcriptions)
- Analyse des relations et représentation cartographique détaillée de leur présence et de leurs relations à travers les réseaux économiques, culturels et sociaux.
- Constitution et organisation des données (parallèlement à l'enrichissement de la Phonothèque prévue dans le programme CRITERR), construction de corrélations (incluant les représentations cartographiques)
- Analyse théorique sur les questions du traitement de l'information concernant des propriétés attachées à des personnes plus qu'à des territoires et sur les problèmes linguistiques posés par les pratiques des agents et la transformation des systèmes (phonologiques et morpho-syntaxiques).
- Phase d'explication et d'application (enseignement, formation en langue des primo-arrivants, aide à la gestion et à la décision des partenaires sociaux et des administrations).

Retombées scientifiques, techniques, socio-économiques et environnementales

Ce projet concerne l'accueil et l'intégration des publics migrants et des populations non sédentaires. Il devrait permettre d'améliorer leur insertion, leur accompagnement, leur formation en fournissant de nouvelles bases à la formation des formateurs du secteur public ou associatif, notamment pour tout ce qui touche à la question linguistique. Il permettra aussi, en fournissant des données fiables et représentatives, d'apporter une aide à la décision politique et à l'évaluation des actions publiques. Du point de vue scientifique, il permettra de nouvelles approches du contact des langues, de ses conséquences sur chacune des langues en présence (émergence de nouvelles variétés : phénomènes de pidginisation, créolisation...) ainsi que sur leur acquisition et transmission. Il permettra de poser le problème de la représentation cartographique de propriétés non territoriales (modélisation des

transferts, banque de données...) et, croisant le point de vue du linguiste et celui du géographe, tout en prenant en compte les pratiques réelles et les représentations des personnes, de reconsidérer des notions comme celles de *territoire* ou de *communauté linguistique*. Projet pionnier, il est destiné à servir de prototype à des enquêtes similaires conduites dans d'autres régions de France ou d'Europe.

III-1- Objectifs et contexte

Parce que chaque fois qu'un individu transmet, apprend ou emploie une langue « il accroît l'utilité de cette langue » (*loyauté linguistique*) aux dépens des autres langues, la mondialisation économique restreint tendanciellement l'offre linguistique au bénéfice de l'anglais. Nombreux sont ceux, linguistes, pédagogues, mais aussi politiques, qui s'interrogent sur la disparition des langues, dont le rythme tend à s'accélérer, ce qui représente la perte d'un patrimoine non reconstituable. C'est, entre autres, dans cette perspective que le Conseil de l'Europe a impulsé un ensemble d'initiatives pour le développement d'une véritable éducation plurilingue et pluriculturelle. Ces actions ont été accompagnées de la commande de matériaux pédagogiques et didactiques (Cadre Européen de Référence des Langues, portfolio européen, etc.). Si ces dispositions vont dans le sens de la construction d'identités plurielles et sont appréciées pour le rôle positif qu'elles ont dans les prescriptions scolaires et universitaires, on constate parallèlement qu'elles profitent essentiellement aux langues officielles des états-nations.

Dans le même temps, les flux de population à l'échelle mondiale et la grande diversité des types de migrants (retraités des pays riches de l'Europe du nord vers l'Europe méridionale, personnels qualifiés et chercheurs en entreprises et grands organismes, réfugiés politiques, gens du voyage, émigrés économiques, etc.) et la variété des pays d'origine et de destination ont pour conséquences de nouveaux contacts humains, l'élargissement du champ de diffusion de certaines langues (langues asiatiques ou africaines parlées en Europe) et de nouveaux contacts linguistiques, impensables il y a deux ou trois décennies, du moins à cette échelle.

De même que l'expansion européenne au 16^e et au 17^e siècle a eu pour conséquence la diffusion des langues d'Europe et l'apparition de nouvelles langues – les créoles –, avec les flux migratoires contemporains apparaissent de nouvelles formes de plurilinguisme qui ont pour conséquence l'émergence de nouvelles variétés linguistiques. Si certaines comme le spanglish des Hispano-Américains ou le chiac (mixte de français et d'anglais, cf. travaux de M.-E. Perrot) parlé au Nouveau-Brunswick (Canada) ont pu donner lieu à des recherches et même être valorisées (publications de livres, émission de radio, presse...), le plus souvent ces embryons de langues nouvelles sont soit totalement méconnus, soit stigmatisés (stigmatisation qui reflète souvent celle qui frappe les locuteurs).

La France, et la région Centre en particulier, est concernée au premier chef par cette problématique (<http://www.oecd.org/home/>). L'arrivée de migrants et l'affaiblissement des variétés dialectales du français (liée à l'allongement des temps de scolarisation et à la rurbanisation) a bouleversé le paysage sociolinguistique au point que certains chiffres, lus sans précaution, pourraient laisser croire que sur le territoire métropolitain, l'arabe a pris la place de l'occitan dans le Midi ou le chinois celle du breton en région parisienne. Dans le même temps, les acteurs sociaux engagés sur le terrain témoignent de l'apparition de formes du français qui leur sont difficilement compréhensibles.

Si la réflexion sur ces thèmes est aujourd'hui avancée aux Etats Unis (hispano-américains) et au Québec où existe une tradition d'enquête langue par langue (cf.

<http://www12.statcan.ca/francais/census01/Products/Analytic/>), en France la prise de conscience est récente et la première initiative publique correspond à l'enquête INED, (cf. www.ined.fr/en/resources_documentation/publications/) initiative relayée par la Délégation Générale à la Langue Française et aux Langues de France (DGLFLF – Ministère de la Culture et de la Communication) et l'ANR (Agence Nationale de la Recherche) qui soutiennent le LLL dans son programme « Langues en Contact à Orléans », l'un des modules associés au projet ESLO2 (Enquête Socio-Linguistique à Orléans n° 2).

Dans la continuité de ces actions, PANGLOSS (**P**rogramme d'**A**nalyse **G**éographique et **L**inguistique à **O**rléans : **S**ciences et **S**ociété) se fixe comme objectifs de répertorier les pratiques linguistiques et culturelles sur l'agglomération d'Orléans, d'analyser les modes d'acquisition, de transmission et de diffusion des différentes langues, de décrire certaines des variétés émergentes (par exemple le peul parlé à Orléans, le français des Voyageurs etc.), de mettre à disposition des échantillons sonores des différentes langues en contact dans un état de préservation (numérisation des données et métadonnées, copies de sauvegarde, double dépôt avec la Phonothèque Nationale, déclaration des formats d'archive, initiative OLAC et utilisation de logiciels Open Source), d'utilisation (autorisation des ayants droit, déclaration CNIL) et d'identification linguistique et sociologique correspondant à l'état le plus récent des connaissances et des exigences en la matière. Au-delà des pratiques en usage en linguistique, l'un des objectifs les plus novateurs de PANGLOSS est l'association avec un laboratoire de géographie reconnu dans ce domaine et qui a déjà conduit une réflexion et abouti à des réalisations en matière de cartographie des pratiques sociales et culturelles.

La cartographie des pratiques linguistiques et culturelles associées à l'usage d'une langue est un défi scientifique du plus haut intérêt dans la compréhension de l'inscription territoriale d'une société. La langue est l'un des éléments essentiels pour identifier une communauté, à la fois dans la reconnaissance interne entre les locuteurs de la langue et dans leur relation aux locuteurs des autres langues. Son inscription dans les territoires permet de mesurer le degré de fragmentation de l'espace social, qu'il importe de saisir dans toute sa complexité pour qu'une politique de la ville aille au-delà des actions sur le foncier et invente des propositions pour obtenir un véritable effet de levier sur le corps social dans le dialogue entre les cultures et la création d'un lien collectif qui transcende les cultures tout en les préservant.

En effet, si la langue crée un espace d'intercommunication qui dessine les échanges à l'échelle d'un quartier, d'une ville, d'un pays, elle peut être aussi un facteur d'isolement social et devenir une source de précarité : difficultés scolaires, difficultés d'insertion professionnelle, problèmes de santé publique, etc. On s'intéresse particulièrement à la concentration relative des populations allophones dans des « enclaves ethniques » qu'il convient de rapprocher d'autres éléments comme la présence de réseaux de chalandise constitués par solidarité culturelle et d'indicateurs de fragmentation sociale du territoire urbain tels que le chômage pour montrer le lien qui peut exister entre telle ou telle communauté et les fractures urbaines. L'intérêt de l'analyse est de contrer les réactions de stigmatisation en ciblant de façon optimale les moyens d'une stratégie destinée à favoriser l'insertion scolaire des enfants, l'insertion professionnelle des adultes ou des actions de promotion de la santé publique.

La représentation cartographique des résultats d'enquête soulève d'intéressantes questions relatives à la distinction entre les espaces de résidence et les quartiers fréquentés quotidiennement. Elle interroge, dans une société de croisement des données, les possibilités de préserver la confidentialité des renseignements individuels, la représentativité de l'échantillon ou encore la mesure du degré de ségrégation spatiale entre différents groupes. Pour répondre à ces questions, l'équipe du CEDETE entend mettre en œuvre ses compétences au travers de :

1. la constitution d'un SIG urbain permettant d'intégrer les résultats de l'enquête aux échelles pertinentes et de les rapprocher d'indicateurs du contexte démographique et social des quartiers, ainsi que d'autres éléments tels que la localisation de certains types de commerce (qualifiés de communautaires) etc.
2. la représentation cartographique des résultats de l'enquête mettant en œuvre des procédés de traitement innovants à l'aide des méthodes de la statistique spatiale comme l'analyse des proximités et des associations et le lissage cartographique qui permettront de traiter les problèmes de représentativité de l'échantillon et de confidentialité des données individuelles.
3. l'analyse des résultats de l'enquête en termes de ségrégation à l'aide des méthodologies (Massey et Denton,1988).

Les résultats attendus doivent constituer un outil d'aide à la décision pour la politique de la ville. Ils constitueront pour les opérateurs, comme l'ACSE ou l'Agence Nationale pour la Rénovation Urbaine, avec lesquels des contacts sont d'ores et déjà établis, des instruments précieux dans l'accomplissement de leurs missions. Ils apporteront de nouvelles perspectives aux différentes formations de la Région (IUFM en formation initiale des maîtres, formation continue et formation de formateurs, différents Masters des domaines linguistique, Sciences humaines, diplôme universitaire « Apprentissage, langues et interculturalité », formation de travailleurs sociaux), ainsi qu'aux associations engagées dans la formation des migrants. C'est par là qu'elles pourront avoir un impact direct ou indirect sur le marché de l'emploi, dans la qualification des opérateurs notamment.

III-2- Description du projet et résultats scientifiques et techniques attendus

PANGLOSS est un projet particulièrement novateur dans son approche des questions linguistiques liées à l'immigration. En effet, si depuis quelques années, l'idée commence à faire son chemin que la France aussi est une terre de multilinguisme, aucune étude à ce jour n'a proposé d'aborder la question en liant à la fois la description externe (géographie linguistique, sociologie du langage) et description interne (analyse des structures phonologiques, morpho-syntaxiques, sémantiques des langues en contacts et des nouvelles variétés linguistiques). Cette approche permettra de surmonter deux limites inhérentes aux projets de recherche centrés sur cette thématique : un commentaire à visée sociologique qui recense l'utilisation des langues en présence en restant circonscrite sur le plan externe des représentations ; une description linguistique qui par manque d'explication du contexte se révèle difficilement exploitable. Elle vise, par l'éclairage de la géographie, à passer d'une présentation statique des données à une figuration dynamique.

Le poids des représentations personnelles et/ou collectives sur toutes les questions qui touchent aux langues et à leur utilisation est tel que toute donnée doit être remise en perspective du point de vue de ces représentations. Celles-ci apparaîtront en particulier au travers de la comparaison entre le « déclaratif » et les pratiques réelles, ce qui implique que les enquêtes seront menées selon trois points de vue :

Le « macro déclaratif » : Enquête à base de questionnaires distribués massivement et sélectivement, recueil d'indications quantitatives et qualitatives auprès d'associations et ou institutions, étant admis que l'absence de données ou l'impossibilité d'en fournir représente *per se* une donnée. On soulignera que les données collectées ne peuvent être interprétées que si elles sont lues à la lumière des résultats obtenus grâce à l'approche « micro déclarative » et à une observation construite des pratiques réelles.

Le « micro déclaratif » : entretiens semi-guidés avec les membres des différentes « communautés linguistiques », en les positionnant à des degrés divers de compétence (médiateurs de la communauté, personnalités du groupe, néo-immigrés, couples « mixtes »...). Au cours de ces entretiens les thèmes abordés tourneront autour de l'utilisation des langues en contact dans les différents types de situation, y compris les accès médiatiques (chaînes par satellite).

Les pratiques réelles : Analyse interne des entretiens semi-directifs ; recueil et analyse de productions orales diverses (conversations, textes...) dans et à propos des différentes langues (y compris le français), avec une part d'observation dans des situations non contrôlées (grandes surfaces, marché, sortie d'école...).

Le réinvestissement, au cours des enquêtes et de l'analyse des données, des acquis méthodologiques et scientifiques de travaux menés par différents membres de l'équipe sur d'autres terrains est un autre défi du projet. Plusieurs chercheurs du LLL sont les auteurs de travaux sur des situations de contacts de langues dans d'autres pays : Afrique lusophone (créoles de Guinée Bissau, de Casamance, de Sao Tomé et Príncipe et du Cap-Vert, variété populaire du portugais en Afrique et au Brésil); Guyane (langues amérindiennes, créole français), Burkina-Faso (langues africaines et français), Canada (chiac du Nouveau-Brunswick). Confronter les résultats de ces travaux avec ceux obtenus lors des enquêtes à Orléans permettra de dépasser l'opposition entre l'universel et le particulier. Les résultats attendus concerneront donc non seulement la situation orléanaise mais seront aussi des éléments pour une analyse à plusieurs niveaux (local, régional, national, international) des phénomènes de différentes natures liés aux contacts de langues : acquisition, transmission familiale (*linguistic loyalty*), diffusion, scripturisation, emprunts lexicaux, alternance codique (code switching), véhicularisation/vernacularisation, pidginisation, créolisation...

Un des grands défis relevé par PANGLOSS est aussi de croiser sur la question des langues le regard du linguiste et celui du géographe. Ainsi sera posé le problème d'une représentation cartographique qui s'attachera surtout aux personnes, à leur trajectoire et à leurs pratiques (modélisation des transferts, banque de données...). S'attachant aux pratiques réelles et aux représentations des personnes, PANGLOSS œuvrera pour la déconstruction de notions comme celles de territoire, de communauté linguistique, ou même d'identité. La principale ambition réside dans le caractère pionnier d'une recherche qui puisse servir de prototype à des enquêtes similaires dans d'autres régions de France ou d'Europe.

Ce projet est une première réalisation correspondant aux orientations proposées par un groupe de laboratoires de l'Université d'Orléans dans la mise en place du Centre de Ressources et d'Innovation sur les TERRitoires en Région Centre, qui constitue le pôle orléanais de l'axe de recherche régional CREVIT « Villes, Territoires, Civilisations » représenté à Tours par le CREVILLES. Il constitue le premier élément d'un rapprochement entre le laboratoire de linguistique LLL et le laboratoire de géographie CEDETE.

Les recherches du LLL incluent une visée anthropologique qui figure en toutes lettres dans le sigle du laboratoire. Cette approche s'est concrétisée dans l'étude des transformations de langues en situation de transfert et de mélange de populations (créoles à base portugaise d'Afrique), dans la relation des langues aux cultures dont elles sont inséparablement solidaires (Encyclopédie des langues de Guyane : palikur, wayana,

wayampi) mais aussi dans la différenciation du français hors de France (Burkina) et en France.

Reprenant une enquête conduite en 1968-1971 par une équipe d'universitaires anglais afin d'obtenir un « portrait sonore » de la ville, le LLL a conçu une enquête comparative dont il réunira les premiers matériaux au cours du premier semestre 2008. A la recherche entreprise il y a quarante ans sous l'intitulé ESLO (Enquête Socio-Linguistique à Orléans, désormais cité comme ESLO1), les linguistes de l'Université d'Orléans se proposent d'apporter le renouvellement d'une enquête du même type, dénommée ESLO2, qui apprécierait les variations et les changements survenus dans la dynamique du changement linguistique. PANGLOSS figure dans ce cadre comme l'un des modules complétant les entretiens sur échantillon représentatif qui constitueront ESLO2, celui qui intégrera à la dynamique du français parlé les transformations et les enrichissements induits par le contact de langues.

Parallèlement, les travaux conduits par le CEDETE en matière de systèmes d'information géographique et de représentation cartographique, en particulier dans la résolution des problèmes que posent des propriétés des agents qui, pour être présentes sur certaines portions de l'espace, n'y sont pas intrinsèquement liées, a fait de la question de la représentation des langues un premier terrain d'expérimentation pour une collaboration qui ira en s'approfondissant. La dimension d'information, du public et des décideurs, et l'aide à la gestion raisonnée des territoires et des politiques urbaines, constituent l'une des finalités de ces programmes.

Quant au troisième laboratoire fondateur du CRITERR, le LEO, il s'associera lors de l'extension d'échelle de cette recherche, en particulier dans la relation entre langue et travail, dans une collaboration qui impliquera le CEREQ et l'équipe TOTEM (Territoire, Organisation, Travail Et Mondialisation) du LEO, laquelle est destinée à réunir les chercheurs en "économie internationale" "économie du travail" et "économie industrielle".

C'est pourquoi il est prévu de loger dans les infrastructures immobilières du CRITERR l'équipe qui se consacrera particulièrement au PANGLOSS, d'abord pour conforter la synergie des équipes, ensuite pour assurer les transferts de savoir et de savoir-faire, enfin pour assurer les relations avec les projets logés dans le Centre au titre du LLL (COSLA régional, Phonothèque, Observatoire des Langues, Encyclopédie des langues de Guyane...), du CEDETE et du LEO.

A l'échelle de la région, PANGLOSS est conçu comme une première proposition destiné à tester la faisabilité d'un projet qui étendrait à terme son périmètre à l'échelle de la région tout entière et la pertinence de collaborations qui sont en construction, l'accord rapide entre les parties sur des objectifs et des méthodes étant du meilleur augure.

Le rapprochement en cours entre une équipe de linguistique de Tours et le LLL participe de cet effort de constitution d'un pôle de référence à l'échelle européenne en région Centre, au même titre que la réunion, en une fédération de laboratoires (FELINCO) d'équipes linguistiques de Poitiers (FORELL), de Tours (Langage et Représentation), de Limoges (CERES) afin de développer des synergies en matière de recherche, de formation (création d'un Master inter-régional) et de transfert de technologies. Cette collaboration régionale est destinée à s'ouvrir aussi au laboratoire de sociologie de Tours qui a fait de l'immigration l'un de ses thèmes de recherche privilégiés.

En termes de développement du potentiel de recherche de la Région Centre, il s'agit de faire émerger, à côté du pôle cosmétique et de grands laboratoires déjà reconnus (sciences de la terre et énergétique à Orléans), des compétences en sciences humaines à même de faire apparaître et de soutenir une activité économique dans des domaines émergents : traitement de l'information et systèmes d'information géographiques, industries de la langue, génie linguistique, traitement automatique des langues. L'extrême faiblesse en particulier des industries de la langue en Région Centre (à la différence, en dehors de Paris, de ce qui s'est passé à Grenoble, Nancy, Aix ou Toulouse) suppose une stratégie volontariste de renforcement d'équipes positionnées dans ce domaine et capables d'attirer, promouvoir ou développer une activité économique en ce domaine.

Dans l'immédiat, le défi auquel répond PANGLOSS correspond à la nécessité de rassembler des compétences afin d'assurer la maîtrise scientifique d'objets d'études qui ont des implications, au-delà des langues, sur l'occupation des territoires, les organisations sociales (cf. l'histoire des migrations) et qui requièrent des méthodes particulières de traitement (banques de données) et d'analyse (cartographie). Une proposition de travail avec l'ANRU (Agence Nationale de Rénovation Urbaine) est en cours d'instruction dans leurs services et c'est dans une telle connaissance des situations et des personnes que se révèle le caractère crucial des études de terrain, dans le passage de la description à l'explication et de l'explication à la prescription par les pouvoirs publics et les administrations.

GUYANE

Encyclopédie de trois peuples amérindiens

Conseil scientifique

Serge Bahuchet, Professeur d'ethnobiologie au Muséum National d'Histoire Naturelle, Paris

Gabriel Bergounioux, Professeur de linguistique à l'Université d'Orléans

Gérard Collomb, anthropologue, chargé de recherche au CNRS, Paris

François Nemo, Professeur de linguistique à l'Université d'Orléans

Dominique Tilkin-Gallois, Professeur d'anthropologie à l'Université de São Paulo, Brésil.

Collection « ENCYCLOPÉDIES »

Soucieux d'articuler travail linguistique, prise en compte des dimensions sociale et historique ainsi que restitution d'une culture dans son intégralité, des spécialistes français et étrangers ont entrepris la réalisation de trois encyclopédies dédiées à trois peuples amérindiens de Guyane, les Palikur, les Wayana, les Wayâpi.

L'originalité de notre projet se définit par son ambition, présenter l'ensemble des savoirs existants chez ces trois peuples et à leur sujet. Pour réaliser cette ambition, nous avons choisi une double voie : d'abord, nos encyclopédies prendront forme dans un ensemble de fascicules thématiques paraissant progressivement et couvrant tous les aspects de la culture, des grands témoignages du passé aux synthèses thématiques d'aujourd'hui ; puis, au terme de chacune des trois séries, les données publiées seront collationnées dans deux ouvrages de synthèse, un dictionnaire culturel et un dictionnaire de langue.

À l'heure où le savoir est littéralement émietté et ce faisant, très largement inaccessible, ce format est une innovation qui nous semble aujourd'hui devenue indispensable : les savoirs n'existent que s'ils ne restent ni méconnus ni isolés, et ils ne vivent véritablement que lorsqu'il est possible à tous de se les approprier. Cette forme de restitution répond à une forte demande sociale. En effet, une longue pratique du terrain et des hommes nous a conduit à la certitude que, conscientes des dangers de 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 désormais aux vertus de la chose écrite.

Chacune de nos trois encyclopédies a pour objet, plus que la description de mots, 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 de l'intérieur, avec pour fil d'Ariane la langue, dans son fonctionnement comme dans ses réalisations. Les ouvrages que nous avons entrepris ne sont pas unilingues. Ils 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.

Nous avons choisi de constituer une collection de publications autonomes à visée globalisante, afin que, 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. Chapeautant notre entreprise, un fascicule zéro intitulé *Langue, milieu et histoire*, offre une présentation de l'ethnie pour chacune des trois encyclopédies, *palikur*, *wayâpi*, *wayana*, et énonce la philosophie de notre démarche.

Françoise Grenand, directrice de la collection

1. PLAN D'ENSEMBLE

Même lorsqu'il leur arrive d'être menées de conserve par certains des auteurs (et bien sûr des éditeurs scientifiques), les trois encyclopédies demeurent des aventures qui se doivent, par respect envers les traits culturels propres aux trois populations, de garder une certaine individualité. Chacune des trois étant profondément ancrée sur la langue, le milieu et l'histoire, l'agencement des connaissances qui fondent la culture revêt une importance capitale.

1.1. Grands domaines des fascicules

Les fascicules se développent autour de plusieurs thèmes majeurs. Il est du ressort des éditeurs scientifiques de veiller à ce que la collection, une fois terminée, les aient fait apparaître sans en omettre. Afin de ne laisser aucune donnée sur le bord du sentier, il a été décidé qu'elle pourrait être amendée ; la décision du découpage et de l'organisation restant fonction de la compétence des chercheurs, la liste qui suit demeure purement indicative.

Domaines majeurs	Thèmes	
• histoire	généalogies histoire des communautés grands témoignages	guerre et paix biographies peuples voisins
• géographie	toponymes spatialisation orientation	topographie astronomie cosmologie
• sciences naturelles	agriculture et tenure foncière faune flore sols	classification indigène couleurs
• technologie culturelle	art et artisanat techniques art musique	jeux parures architecture et habitat aspects de la modernité
• vie sociale	personne entre corps et âmes maladie et médecine nomenclature de parenté étapes de la vie organisation sociale droit coutumier et morale chefferie	chamanisme cuisine et art culinaire division sexuelle du travail aspects de la modernité : école salarial religions importées
• littérature orale	danses chantées chants poésie	mythes récits proverbes et dictons
• langue	questions de phonologie questions de grammaire questions de sémantique questions de pragmatique usages sociaux du langage	expression de la modernité langues reliques langues secrètes différentes écritures représentation de la langue
• dictionnaires	dictionnaire fondamental bilingue	Vaste dictionnaire encyclopédique final

• littérature enfantine	faune flore	techniques contes...
-------------------------	-------------------	---------------------------

1.2. Cohérence de l'encyclopédie

Chaque encyclopédie trouve sa cohérence dans le déroulement global des grandes sections suivantes en deux étapes :

- 1• la présentation générale,
 - les études,
 - la bibliothèque,
 - des publications pour la jeunesse,
 - un dictionnaire fondamental,
- 2• le grand dictionnaire de langue,
 - les index.

• *La présentation générale*

Chacune des trois encyclopédies s'ouvre par une présentation générale, qui en constitue le fascicule zéro. Elle propose la philosophie et les orientations de notre travail, le plan général de l'encyclopédie, la liste ouverte des collaborateurs potentiels et, de fait, le précis des conseils aux auteurs. Elle contient aussi une présentation de l'ethnie concernée, l'état de l'art, ainsi qu'une bibliographie la plus exhaustive possible.

• *Les études*

Synthèses nouvelles spécialement rédigées pour l'encyclopédie, leur nombre n'a pas été arrêté, mais au contraire volontairement laissé ouvert, le but étant de couvrir les domaines majeurs précédemment définis. Les fascicules peuvent être différents dans leur plan, leur ordre de parution et leur nombre final, mais ils se conforment à l'esprit donné par l'éditeur scientifique, après agrément du conseil éditorial. En nombre variable selon les ethnies, ils constituent le cœur de l'encyclopédie. Chacun est une étude scientifique classique, avec de nombreux exemples numérotés, tentant, à chaque fois, la synthèse des connaissances dans le domaine traité. Il comporte des dessins, des cartes et des photographies chaque fois que le thème les impose. Il porte un titre en français, et, dans la mesure du possible, un équivalent dans la langue vernaculaire. Selon nécessité, certains fascicules comportent un index thématique français, ainsi que deux lexiques (français-langue vernaculaire, langue vernaculaire-français).

• *La bibliothèque*

Nous appelons *bibliothèque* l'ensemble des textes de littérature orale produit par chacun des trois peuples et les textes d'archives produits par ceux qui les ont visités. Le fait qu'elle rassemble littérature orale autochtone et sources écrites allogènes lui confère un statut original et singulier auquel nous tenons. Les textes ici rassemblés s'avèrent être les deux versants de la connaissance sur les peuples dont nous parlons. Cette littérature existe, mais elle est souvent soit dispersée, soit non publiée, enfouie dans des archives ou bibliothèques privées, quelquefois même, non encore collectée. Nous nous proposons de rassembler des textes anciens, des inédits ou des traductions, de les publier et d'ainsi mettre pour la première fois à la disposition tant des membres de la communauté autochtone que d'un public élargi l'ensemble des textes connus, recueillis ou non, découverts ou non par les membres de l'équipe. Chaque fois que cela est possible, est présenté le texte dans la langue vernaculaire et dans une traduction française. Dans le cas de textes d'archives, des présentations et des notes (historiques, critiques, linguistiques...) sont indispensables. Les pièces de ce corpus uniquement disponibles en d'autres langues (en l'occurrence le portugais, l'anglais, l'allemand ou le hollandais) sont données en traduction française. Une des tâches des éditeurs scientifiques sera de veiller autant à n'oublier aucun de ces documents qu'à éviter toute redondance avec d'autres fascicules, et l'on pense ici plus particulièrement à la mythologie.

- *Les publications pour la jeunesse*

La jeunesse des trois ethnies ne cesse de gronder devant le manque criant de matériaux didactiques à son usage, cette demande sociale croissante allant de pair avec la conscientisation politique des jeunes générations.

Ainsi, les jeunes Wayâpi par exemple nous ont demandé de rédiger, dans un français usant de mots simples, mais ce pourrait aussi être un texte bilingue, une histoire de leur peuple pour les enfants, ainsi que des fascicules sur les savoirs naturalistes de leurs parents.

Pour les jeunes Wayana, une autre idée a vu le jour, celle d'ateliers d'histoire, en concertation avec l'école : des enquêtes seraient menées auprès des anciens par les jeunes adolescents sur les récits historiques, dont les résultats seraient ensuite comparés, compilés, amendés et illustrés par les enfants, avant d'être publiés dans l'encyclopédie, sous forme bilingue.

Nul doute que dans ce vaste domaine de la littérature enfantine, d'autres projets jailliront.

- *Le dictionnaire fondamental*

À mi-parcours, faisant le lien d'une part entre le grand dictionnaire de langue final et les index, d'autre part entre les études et la bibliothèque, nous avons prévu la parution d'un dictionnaire fondamental bilingue colligeant les mots les plus usuels à partir de leur fréquence relevée dans les enregistrements et les transcriptions.

- *Le grand dictionnaire de langue*

Nous l'ambitionnons le plus complet possible. Véritable encyclopédie culturelle intégrée, par le choix de ses entrées, il reprend de façon synthétique l'ensemble des notions, des concepts, abordés dans les différents fascicules. Au fur et à mesure de la parution des fascicules, le chef du projet aura, entre autres missions, la tâche exaltante et périlleuse d'y puiser la matière vive, et ainsi d'incrémenter tout au long de l'aventure les trois bases de données lexicographiques et dictionnaires. Lorsque toutes les études seront terminées, viendra le temps de publier, sous la direction d'un maître d'œuvre spécifique, un dictionnaire culturel pour chacune des trois langues.

Tous les mots de la langue apparaîtront selon l'ordre alphabétique strict. On aura écarté aussi bien l'enchaînement logique ou analogique des entrées, tenté dans le *Dictionnaire wayâpi* (F. Grenand, 1989), que l'ordre phonologique adopté par les rédacteurs de l'*Encyclopédie des Pygmées Aka* (Thomas & Bahuchet, 1983 et suiv.) : le premier ne reflète bien souvent que la seule logique du compilateur au moment où il rédige ; le second est fondé sur une structure certes cohérente mais fermée ; tous deux complexifient par trop la lecture du néophyte. Pour arbitraire qu'il soit, l'ordre alphabétique, désormais connu de tous, demeure le plus simple à manier.

Afin de ne pas rester à l'état de squelette, les entrées seront agrémentées d'autant d'exemples qu'il sera jugé bon pour expliciter les différents sens de chaque mot.

Le dictionnaire sera accompagné d'un copieux lexique français-langue vernaculaire, établi avec un soin particulier.

- *Les index*

Chacune des encyclopédies se terminera par un volume d'index servant d'entrée et de repère au questionnement extérieur. Plusieurs index sont nécessaires. Tous se doivent d'être parfaits.

- *Index français raisonné* : les grands champs, en gras, découpés en sous-thèmes, doivent y apparaître, avec mention des volumes et toutes les pages d'occurrence.

- *Index des noms propres* : aussi indispensable que le précédent, il comprendra tous les toponymes, noms de cours d'eau, héros mythologiques ou historiques...

- *Table des illustrations* : pareillement éditée en fin de série, cette table rappellera pour l'ensemble des volumes parus, toutes les occurrences de dessins, photographies, gravures, cartes, plans, planches illustrées...

- *Bibliographie générale* : chaque fascicule comprend sa propre bibliographie. Le fascicule zéro, celui qui annonce tous les autres, en comprend également une. Pour autant, une bibliographie générale finale nous a semblé pertinente, car on la transformera en véritable somme bibliographique de l'ethnie considérée. Placée en fin d'édition de chacune des trois encyclopédies, elle permettra la nécessaire mise à jour, incluant la moisson originale (en particulier documents d'archives ou textes miméographiés) que chaque contributeur n'aura pas manqué de faire.

2. ENCYCLOPEDIE ILLUSTRÉE

Une encyclopédie se doit de ménager une place importante au message iconique (Alain Rey, 1982 : 44). C'est ainsi que, dès le départ, nous avons choisi d'accompagner le projet encyclopédique d'une base de données informatisée¹. Le projet consiste à rassembler et conserver tout document iconographique ou sonore touchant les trois peuples. Cette base de données fait suite à un inventaire général - encore en cours - de fonds iconographiques et muséographiques divers, publics ou privés, dont ceux de l'équipe. Les nombreuses illustrations des fascicules édités y ont été puisées. Chaque fois que cela s'est avéré nécessaire, des autorisations de reproduction furent demandées : pour la faune, la flore, des reproductions d'objets, des gravures anciennes...

2. 1. Photographies, gravures

Le noir et blanc a été, pour d'évidentes raisons de coût, privilégié sur la couleur, laquelle sera réservée à des cas très particuliers (parures de plumes, bibliothèque enfantine...).

Répondant en cela à une demande autochtone maintes fois réitérée, nous avons été particulièrement vigilants à rechercher les *portraits* (gravés ou photographiés) des grands hommes des trois peuples. On peut d'ores et déjà dire que les Wayana, grâce à Coudreau puis Hurault, sont les mieux servis. Pour les Wayãpi, on a aussi cherché à illustrer les lieux de la grande migration du XVIII^e siècle. Pour les Palikur, ceux du berceau de l'ethnie en Amapá.

2.2. Dessins

On y a eu recours aussi souvent que nécessaire. Pour l'artisanat, on a fait dessiner au trait tous les objets (*de visu* ou d'après photographie), si possible en les resituant, par comparaison avec d'autres, dans l'espace et le temps ; ainsi sont visualisées les influences des ethnies voisines (espace) et les transformations internes (temps).

2.3. Bandes dessinées

Pour le déroulement d'un savoir faire, d'une technique, d'un tour de main, on a utilisé la technique de la BD en quelques images : par exemple monter une poterie au colombin ; décrocher un balai fructifère du haut d'un palmier.

2.4. Planches

¹ Elle est constituée suivant le cahier des charges défini par les instances internationales comme OLAC, à partir de logiciels « *Open Source* » et avec contrôle de la diffusion afin de préserver les intérêts des ayants droit. La normalisation des fichiers, des transcriptions et des principes de catalogue permettra une interrogation ouverte et une évolution des archives pour une utilisation large et évolutive.

La modernité, dont on a dit qu'elle occupe une place importante, est elle aussi abordée par le dessin. Par exemple, des planches détaillent les parties des objets modernes simples (une bouteille, un moteur hors-bord) ou complexes (la ville, l'école), avec toutes les parties ou sous-parties nommées en langue vernaculaire.

2.5. Cartes

Sans avoir la prétention de doter les trois encyclopédies d'un atlas, nous avons accordé aux cartes un soin particulier ; elles ont pour but de retracer les chemins de migration, les circuits de commerce coloniaux, de délimiter les terroirs, d'en marquer l'évolution, de figurer les ethnies voisines. Dans le contexte particulier de la Bibliothèque, les cartes anciennes sont aussi incluse.

3. CALENDRIER DE PUBLICATION

3.1. Volumes parus

- Wayana 1 2007. Chapuis Jean. *L'ultime fleur*, ekulunpi tihmelë, *essai d'ethnosociogenèse wayana*, PUO, Orléans.
- Palikur 1 2008. Curt Nimuendaju. Traduction de *Les Indiens Palikur et leurs voisins*, présentée par Pierre Grenand.
- Wayana 0 2009. Eliane Camargo & Pierre Grenand. *Une encyclopédie, un peuple*.
- Wayãpi 0 2009. Pierre Grenand & Dominique Tilkin-Gallois. *Une encyclopédie, un peuple*.
- Palikur 0 2009. Antonia Cristinoi, Pierre Grenand & Artionka Kapiberibe. *Une encyclopédie, un peuple*.

3.4. Volumes en préparation

- Palikur 2 2010. Gabriel Bergounioux. *Phonologie du palikur, avec proposition de transcription*.
- Wayãpi 1 2011. Françoise Grenand. *Histoires dans la forêt, contes pour enfants*, édition bilingue.
- Wayãpi 2 2011. Pierre Grenand & Françoise Grenand. *Ethnobotanique*.

Au total, une quarantaine de volumes prévus sur douze ans.

Séminaire du GORDF

2007-2008

- 17 octobre
« Acquisition de la liaison et mise en place de structures syntaxiques dans le cadre du groupe nominal chez les enfants entre 2 et 6 ans », Céline Dugua, MCF, université d'Orléans, LLL
- 23 novembre
« Penser les techniques intellectuelles-scolaires : entre écriture, école et familles », Mathias Millet, MCF, université de Poitiers, SACO (Savoirs, cognition et rapports sociaux)
- 19 décembre
« Une langue de référence pour entrer dans la lecture », Silvia Lucchini, UCL (université catholique de Louvain-La-Neuve) et université de Liège, Belgique
- 16 janvier
« De quelques fondamentaux de la didactique du français », Emile Genouvrier, Professeur émérite, université de Tours
- 6 février
« Présentation de la méthode de lecture "Bulle" (Bordas) par ses auteurs », Patricia Bucheton, MDPH Orléans, GORDF, Marie-Aude Murail, écrivain pour la jeunesse et Christine Thiéblemont, PE
- 12 mars
« Du rapport à l'écriture à la compétence scripturale : apports et limites du concept de rapport à l'écrit », Suzanne Chartrand, université de Laval, Québec
- 2 avril
« Le rapport à l'écriture et l'apprentissage des formes discursives », Christine Barré-De Miniac, Professeur des universités, IUFM de Grenoble, Lidilem
- 14 mai
« L'éveil aux langues en classe ordinaire », Catherine Brumelot, CASENAV Orléans, GORDF
- 11 juin
« La construction de la compétence d'Interaction Orale (IO) en anglais de spécialité par le TCMO (Travail Collaboratif Médiatisé par Ordinateur) », Cédric Sarré, université d'Orléans, LLL-GORDF

2008-2009

- 22 octobre
« Normes discursives et rapport à l'écrit au début du parcours universitaire : entre implicite et conscientisation », Catherine Deschepper et Francine Thyryon, UCL, Belgique
- 26 novembre
« Didactique du français et image mobile : l'analyse sémio-pragmatique de l'offre éditoriale actuelle pour l'enseignement secondaire », Philippe Bourdier, IUFM Centre Val de Loire LLL-GORDF

▪ 17 décembre : « Les reformulations : des outils linguistiques favorisant l'ajustement et l'inter-compréhension », Blandine Pennec, université de Perpignan

▪ 14 janvier
« Typologie des reformulations dans les conversations de jeunes libanais francophones », Loyal Kanaan, doctorante (LLL- GORDF)

▪ 18 février
« L'écriture collaborative à distance », Jean-Claude Bertin, université du Havre

▪ 18 mars
« L'intérêt didactique de la variation dans les textes littéraires francophones », Diane Schwob, (LLL- GORDF)

▪ 8 avril
« La question du sujet dans les textes prescripteurs de l'école primaire », Marie-France Bishop (IUFM de Versailles, Théodiles)

▪ 6 mai
« La stéréotypie un outil d'enseignement et d'apprentissage de l'écriture ? A quelles conditions ? », Bernadette Kervyn, UCL Belgique et université de Bordeaux 1.

2009-2010

▪ 14 octobre
« Apprendre à lire et à écrire des énoncés de problèmes mathématiques », Annie Camenisch et Serge Petit, IUFM d'Alsace, LilPa

▪ 18 novembre
« Les écritures parentales : histoire, pratiques et enjeux », Véronique Francis, IUFM Centre, Val de Loire, CREF

▪ 16 décembre
Thème : le développement du langage en maternelle, Véronique Boiron, IUFM d'Aquitaine

▪ 6 janvier
Thème : la didactique de la littérature en Europe ou une didactique intégrée du français et des langues, Dominique Ulma, IUFM de Lyon, Dynadiv et LLL-GORDF

▪ 17 mars
« Orthographe : nouveaux enjeux didactiques », Danièle Cogis, IUFM de Paris, MoDyCo UMR 7114

▪ 21 avril
« En amont du brouillon rédigé : nouveaux regards sur la planification du projet d'écriture de l'école au lycée », Olivier Lumbroso, IUFM de Bretagne, CENTRE ZOLA - ITEM / CNRS UMR 8132

PROHEMIO : Collation d'ouvrages pour le relevé des figements

Ouvrages traduits dans lesquels les expressions figées ont été récupérées à ce jour (juin 2010) pour la constitution du corpus du *Dictionnaire idéologique comparatif des expressions figées espagnoles et françaises*.

Camilo José Cela : *La familia de Pascual Duarte* (*La famille de Pascal Duarte*, Seuil, Paris, 1997. Tr. Jean Viet)

Camilo José Cela : *La colmena* (*La ruche*, Gallimard, Paris, 2003. Tr. Henri LP Astor)

Miguel Delibes : *Cinco horas con Mario* (*Cinq heures avec Mario*, Verdier, Lagrasse, 2010. Tr. D. Blanc)

Miguel Delibes : *El camino* (*Le chemin*, Gallimard, Paris, 1959. Tr. Maurice Edgar Coindreau)

Juan Marsé : *El amante bilingüe* (*L'amant bilingue*, Christian Bourgois, Paris, 1996. Tr. J.-M. Saint Lu)

Carmen Martín Gaité : *Nubosidad variable* (*Passages nuageux*, Flammarion, Paris, 1995. Tr. Claude Bleton)

Carmen Martín Gaité : *Lo raro es vivir* (*Drôle de vie la vie*, Flammarion, Paris, 1999. Tr. Claude Bleton)

Eduardo Mendicutti : *El palomo cojo* (*Le pigeon boîteux*, Christian Bourgois, Paris, 1995. Tr. D. Laroutis)

Eduardo Mendicutti : *Una mala noche la tiene cualquiera* (*Une sale nuit*, Ch. Bourgois, Paris, 1992. Tr. Denise Laroutis)

Eduardo Mendoza : *Sin noticias de Gurb* (*Sans nouvelles de Gurb*, Seuil, Paris, 2006. Tr. François Maspero)

Eduardo Mendoza : *La aventura del tocador de señoras* (*L'artiste des dames*, Seuil, Paris, 2002. Tr. F. Maspero)

Javier Tomeo : *Diálogo en re mayor* (*Dialogue en Ré majeur*, Editions de la Mauvaise Graine, Lyon, 2001. Tr. Mor Gaye)

Javier Tomeo : *Amado monstruo* (*Monstre aimé*, Christian Bourgois, Paris, 2007. Tr. Denise Laroutis)

Javier Tomeo : *La máquina voladora* (*La machine volante*, Ch. Bourgois, Paris, 2004. Tr. Denise Laroutis)

Javier Tomeo : *El cazador de leones* (*Le chasseur de lions*, Ch. Bourgois, Paris, 1990. Tr. Denise Laroutis)

A ces ouvrages s'ajoute la quarantaine de titres constituant l'œuvre de Javier Tomeo (environ 3000 pages). Les expressions figées y ont également été recensées mais uniquement dans la version originale (tous les romans n'ont pas été traduits).

CAHIERS DU PROHEMIO

***Cahiers du P.R.O.H.E.M.I.O.*, n° 1, 1996, 169 pp.**

"Oralité, histoire, écriture : réalisations et perspectives".

Actes de la Journée d'Etudes du 15 décembre 1995.

Université d'Orléans.

Edition préparée par Angel Iglesias Ovejero avec la collaboration de Cécile Iglesias, Monique Güell, Françoise Giraud.

***Cahiers du P.R.O.H.E.M.I.O.*, n° 2, 1998, 318 pp.**

"Usages de la formule".

Actes du II Colloque International (21-22 novembre 1997).

Université d'Orléans.

Edition préparée par d'Angel Iglesias Ovejero avec la collaboration de Cécile Iglesias, Monique Güell, Françoise Giraud, Daniel Vassivière.

***Cahiers du P.R.O.H.E.M.I.O.*, n° 3, 2000, 453 pp.**

"Expressions figées : idiomaticité, traduction".

Actes du III Colloque International (26-27 novembre 1999).

Université d'Orléans.

Edition préparée par Angel Iglesias Ovejero avec la collaboration de Denise Boyer, Lidio J. Fernandez, Françoise Giraud, Monique Güell.

***Cahiers du P.R.O.H.E.M.I.O.*, n° 4, 2002, 379 pp.**

"Idiomaticité, motivation et onomastique : la fabulation étymologique"

Actes du IV Colloque International (22-23-24 novembre 2001).

Université d'Orléans.

Edition préparée par Angel Iglesias Ovejero avec la collaboration de Denise Boyer, Lidio J. Fernandez, Françoise Giraud, Monique Güell.

***Cahiers du P.R.O.H.E.M.I.O.*, n° 5, 2004, 288 pp.**

***Cahiers du P.R.O.H.E.M.I.O.*, n° 6, 2004, 384 pp.**

"Formas de Vida y Cultura Tradicional en El Rebollar y otras partes".

Actes des I Journées Internationales d'Etudes de *El Rebollar*.

V colloque P.R.O.H.E.M.I.O. (23-27 juillet 2003).

El Rebollar (Salamanque).

Edition préparée par Angel Iglesias Ovejero avec la collaboration de Lidio J. Fernandez, Françoise Giraud, Monique Güell.

CAHIERS DU PROHEMIO 7 & 8

Interferencias lingüísticas y culturales en El Rebollar y otras partes

(Actas de las II Jornadas de Estudio de El Rebollar)

VI Coloquio del P.R.O.H.E.M.I.O.

22-25 de julio de 2005

ÍNDICE GENERAL

I Volumen (*Cahiers du P.R.O.H.E.M.I.O.*, VII)

Presentación 15

Mapa de El Rebollar 21

**I. Interferencias lingüísticas:
analogías y contrastes entre las hablas
de El Rebollar y de otras partes..... 25**

Miguel Á. REBOLLO TORÍO,
*Unidad y divergencias lingüísticas en Extremadura
y sur de Salamanca* 27

Ángel IGLESIAS OVEJERO,
*Interferencias lingüísticas y culturales en El Rebollar:
algunas propuestas para la transcripción gráfica
de conversaciones, relatos y testimonios orales
en la modalidad rebollana* 43

Ángel IGLESIAS OVEJERO,
*Analogías y contrastes entre las hablas de El Rebollar
y las extremeñas* 75

M^a Luisa MONTERO CURIEL,
*Cambios fonéticos en Castilla y León:
del ALPI al ALCL* 101

Pilar MONTERO CURIEL,
*El léxico extremeño y salmantino en el Diccionario
de Uso del Español de María Moliner* 115

Ana SERRADILLA CASTAÑO,
*Morfología y sintaxis extremeñas en el norte de
Huelva. Fronteras lingüísticas vs. fronteras políticas ...* 131

M^a Victoria NAVAS SÁNCHEZ-ÉLEZ,
*El léxico de las ocupaciones, influencia de dos áreas
lingüísticas en contacto:
Barancos (Portugal) y Encinasola (Huelva)* 149

Luc TORRES,

Contribución para la elaboración de un Glosario de las voces y modismos de los dialectos y hablas leoneses (leonesismos, montañesismos, asturianismos, salmantinismos, bablas de Extremadura y El Rebollar) presentes en La Pícara Justina) 173

**II. Interferencias lingüísticas:
traducción, bilingüismo, extranjerismos 193**

Josep GUIA, Maria CONCA,
Análisis contrastivo del frásico de Tirant lo Blanc (València, 1490; Barcelona, 1497) y sus traducciones castellana (Valladolid, 1511) e italiana (Venezia, 1538, 1566, 1611). Entre el calco, la equivalencia y la creación 195

Ramon LLADO, Xus UGARTE,
Multilingüismo y diglosia en Rabelais: lo que las traducciones revelan 217

José ALONSO PASCUAL,
Conflictos, clases, ventajas mitos y efectos adicionales del bilingüismo..... 229

Sylvie FOURNIÉ-CHABOCHE,
Las palabras españolas de la lengua francesa: aspectos formales, semánticos e históricos de la transformación cultural 261

Miguel Ángel GARCÍA PEINADO,
La proliferación inadecuada de coetillas y latiguillos en el español actual (galicismos y anglicismos)..... 281

José Antonio VICENTE LOZANO,
Una lengua de TBO: problemas traductológicos con "Mortadelo y Filemón" 293

II Volumen (Cahiers du P.R.O.H.E.M.I.O., VIII)

**III. Interferencias entre cultura oral y escrita:
fijación léxica.
Fraseología, paremiología, onomástica 335**

Francisco RUIZ GIRELA,
Los refranes: testigos de excepción del contacto entre culturas..... 337

Maria CONCA, Silvia PINTER,
Analogies i contrasts entre unitats fraseològiques antropolímiques de llengües romàniques 351

Ángel IGLESIAS OVEJERO, Françoise GIRAUD,
Cuán necios y cornudos son los otros en

El Rebollar y otras partes:
Portuguesis, Serranus y Farinatus
en el refranero xenófobo rebollano,
extremeño y castellano 363

IV. Interferencias literarias y tradición popular:
Motivos tradicionales en la literatura culta.
Autores regionales 401

José RAMIREZ DEL RIO,
Códigos semióticos e interferencias culturales en
los textos castellanos medievales..... 403

María del Pilar COUCEIRO,
Tradicón grecolatina en la obra
de Rosalía de Castro..... 419

Isabel VÁZQUEZ DE CASTRO,
De la ‘tragedia’ a la ‘trigedia’:
interferencia lingüística
y creación de una estética en Los cuernos de
don Friolera de Valle-Inclán 429

Cécile IGLESIAS,
Canciones infantiles revisitadas:
las nanas en la lírica
de Federico García Lorca 443

José Luis PUERTO,
J. M^a. Gabriel y Galán:
tradición clásica y moderna 467

Pilar MAGADÁN CHAO,
Cuatro villancicos de Juan del Encina:
formas populares que sobreviven
en algunas comarcas salmantinas 483

José Ramón CID,
La gaita salamanquina
y el tamboril tradicionales,
en relación a la obra de Juan del Encina 507

Alfonso SAURA,
Vicente Medina, poeta social y
reivindicativo 523

V. Analogías y contrastes en la historia
y el arte iconográfico.
Constitución de la frontera. Figuras históricas locales.
Arte pastoril. Cine y realidad social.
Simbolismo cromático y política 553

Francisco Javier MORALES PAINO,

*Formación y evolución histórica de
la raya luso castellana:
Convivencia y conflictividad social
en torno a las regiones del Alto Coa y
El Rebollar miróbricense (ss. XII-XV) 555*

Emilio GARCÍA LOZANO,
*Fray Martín de Robleda,
primer obispo de Chile (1513-1561) 587*

José Ramón CID,
*Don Serafín Tella Gallego, robledano ilustre
y destacado arqueólogo 611*

Carlos GARCÍA MEDINA,
Arte pastoril del Rebollar 631

Lidio. J. FERNÁNDEZ RODRÍGUEZ,
*Estudio bibliográfico e histórico del documental
sobre Las Hurdes Tierra sin pan
de Luis Buñuel en 1933 645*

James DURNERIN, Catherine VIENET,
Colores y política 675

**VI. Tradición literaria y modalidad de habla rebollana.
Testimonios. Léxico. Romancero. Textos 685**

José Benito MATEOS PASCUAL,
*Razones (razonadas y de peso) para la defensa
del habla de El Rebollar 687*

Yolanda GARCÍA VICENTE,
*Flora rebollana. Plantas útiles y venenosas.
Sus nombres, características y usos 695*

Gabriel CALVO,
*Anotaciones a un romance geográfico
tradicional salmantino 715*

Pedro Javier CRUZ SÁNCHEZ,
*El cuaderno de rezos de Antonio Franco,
de Espino de la Orbada (Salamanca), 1869:
notas a propósito del “Romance del Arado” 733*

Ángel IGLESIAS OVEJERO, Françoise GIRAUD,
*Tres relatos orales rebollanos:
El lobu y los pastoris, Los polvus de las cocas
y La Muerti 757*

PROHEMIO 9 & 10

Memoria histórica - Historia oral : guerra, postguerra y emigración en El Rebollar y otras partes

III Jornadas Internacionales de Estudio de El Rebollar
VII COLOQUIO DEL P.R.O.H.E.M.I.O.
Curso extraordinario de la Universidad de Salamanca
20-22 de julio de 2007

ÍNDICE GENERAL

I Volumen (*Cahiers du P.R.O.H.E.M.I.O.*, IX)

Presentación 17

I. Historia oral y escrita 23

Fernando SÁNCHEZ MARROYO,
*La Guerra Civil. Algunos libros significativos
y ciertas cuestiones pendientes* 25

Julián CHAVES PALACIOS,
*Represión en la Guerra Civil y movimiento guerrillero
de posguerra en la Sierra de Gata y Salamanca* 37

Santiago LÓPEZ GARCÍA, Severiano DELGADO CRUZ,
*Terror de Estado en la retaguardia:
la represión en Castilla y León
en los primeros meses de la Guerra Civil* 87

Ángel IGLESIAS OVEJERO,
*Archivos vivientes: las víctimas del terror de 1936
en El Rebollar y pueblos aledaños* 101

José Luis HERNÁNDEZ SÁNCHEZ, Miguel LIRIO DÍAZ,
*Notas sobre la presencia militar del Tercer Reich
en Salamanca a través de las Causas del
Archivo Militar de Ferrol* 203

Julio FERNÁNDEZ GARCÍA,
Las cárceles del Franquismo..... 215

Floren DIMAS BALSALOBRE,
*Consejo de guerra:
la máquina asesina del franquismo* 233

M^a Luz de PRADO HERRERA,
*Auxilio Social en Salamanca durante la guerra civil:
Organización y desarrollo* 253

Cayetano IBARRA BARROSO,
Recuperación de la Memoria Histórica en Extremadura:

<i>Investigación, localización y exhumación de fosas</i>	271
José Ignacio MONTEAGUDO, Manuel MATELLÁN, <i>Testimonios de la guerra, posguerra y emigración en el Archivo de la Escritura Popular del Bajo Duero</i>	285
Rafael MORALES RUIZ, <i>Entrevistando a mayores. La experiencia del Taller de Historia Oral del Archivo Municipal de Córdoba</i>	303
II. Historia Oral – Testimonios	337
Mar GONZÁLEZ DE LA PEÑA, <i>Interrelación de la historia oral y los archivos en la búsqueda de un desaparecido en la Guerra Civil española</i>	339
Matilde GARZÓN RUIPÉREZ, <i>La Guerra Civil, El Golpe de Estado y la represión franquista a través de mi experiencia</i>	347
Manuel CORRAL BACIERO, <i>Tirar del hilo: La Guerra Civil 1936-39 en Peñaranda de Bracamonte</i>	359
Luis CALVO RENGEL, <i>La represión fascista en Salamanca</i>	371
Isabel MATEOS MATEOS, <i>Hechos ocurridos en los municipios de Robleda y Boadilla en los días 13 y 14 de agosto de 1936 y posteriores gestiones</i>	387
José ALONSO PASCUAL, <i>...Y pudieron contarlos</i>	393
II Volumen (Cahiers du P.R.O.H.E.M.I.O., X)	
II. Historia Oral – Testimonios (cont.)	421
Maria CONCA, Josep GUIA, <i>Muertos y represaliados republicanos en Beneixama (1939-45)</i>	423
Alfonso SAURA SÁNCHEZ, <i>La represión de los vencedores en la Casa del Pueblo del LLano del Beal (Murcia)</i>	443
Luc TORRES, <i>Historia, intrahistoria y visión regionalista de la Guerra Civil: Los días del Alzamiento en Navarra (18 de julio y días próximos)</i>	461
<i>Testimonios de Robleda 1973, 1976 y El Payo 1973</i>	

transcritos y anotados por
Ángel IGLESIAS OVEJERO 473

III. Postguerra y emigración en su contexto 551

Emiliana LEÓN LÓPEZ,
*Españoles en Francia. Emigración y exilio
en La Voulte* 553

Carlos GARCÍA MEDINA,
La cocina del hambre 569

Gabriel CALVO,
*La postguerra: racionamiento, estraperlo
y prostitución* 589

Beatriz SÁNCHEZ VALDELVIRA,
*La arquitectura en El Rebollar: el caso de Robleda
en el contexto de la postguerra y la emigración* 615

Yolanda GARCÍA VICENTE,
*Remedios caseros hechos con plantas
medicinales en la postguerra* 625

IV. Visión de la Guerra en la literatura y el arte .. 645

María Pilar COUCEIRO,
*La guerra en los Cancioneros populares:
Damaso Ledesma, Bonifacio Gil,
Manuel García Matos* 647

Cécile IGLESIAS,
*Los cantares de la guerra en el repertorio tradicional:
¿memoria u olvido de una poesía circunstancial?* 663

José Luis PUERTO,
*La perspectiva de los derrotados
(arte, música, literatura)* 697

Miguel Á. REBOLLO TORÍO,
La guerra en Baroja 713

Ángeles GARCÍA CALDERÓN,
*Lorca y la Guerra Civil en la poesía canadiense:
traducción de poemas al castellano* 729

Carmen BALBUENA TOREZANO,
*La guerra civil española en la literatura alemana
y su traducción al castellano* 753

André Michel FEUGAIN,
*Análisis contrastivo de los carteles
de propaganda durante la Guerra Civil* 765

James DURNERIN, <i>El rescate de la memoria en Maquis</i> <i>de Alfons Cervera</i>	785
Felipe APARICIO NEVADO, <i>Una guerra cinagética</i>	799
Natalie NOYARET, <i>La memoria de la guerra en</i> <i>Soldados de Salamina y</i> <i>La velocidad de la luz de Javier Cercas</i>	811
Isabel VÁZQUEZ DE CASTRO, <i>Memoria y escritura: estrategias de la transmisión</i> <i>de una historia en La voz dormida de Dulce Chacón,</i> <i>Los rojos de ultramar de Jordi Soler y Maquis</i> <i>de Alfons Cervera</i>	829
Lidio J. FERNÁNDEZ RODRÍGUEZ, <i>El maquis antifranquista en la posguerra</i> <i>de Cantabria y su tratamiento en el cine</i>	845

CAHIERS DU PROHEMIO 11

Historia, Habla y Cultura Tradicional

IV Jornadas Internacionales de Estudio de El Rebollar VIII COLOQUIO DEL P.R.O.H.E.M.I.O. 17-19 DE JULIO DE 2009

ÍNDICE

Cahiers du P.R.O.H.E.M.I.O., XI

Presentación	11
I. Historia reciente y testimonios	19
Ricardo ROBLEDO, <i>Reforma agraria y represión</i>	21
Luis CASTRO BERROJO, <i>El embalse del Águeda, los regadíos y la colonización: una reforma técnica frustrada</i>	37
Maria CONCA, Josep GUIA, <i>Anàlisi d'un sumariíssim d'urgència: el núm 7545 d'Alacant, contra Vicent Conca Amorós, alcalde de Beneixama</i>	55
Franciso Javier MORALES PAÍNO, <i>Historia grabada en mis recuerdos. Memoria viva de una supervivencia. "Puchero Zarceño"</i>	95
Luc TORRES, <i>Cuadernos de un desertor del ejército de Franco. Itinerario de un navarro republicano</i>	115
Miguel LIRIO DIAZ, <i>Apuntes sobre el estudio de Fuentes de Oñoro y su Aduana en las postrimerías de la II República y Guerra Civil</i>	171
Ángel IGLESIAS OVEJERO, <i>Ensayo de cronología del alzamiento militar terror y represión de 1936 a 1946 en el Alto Águeda</i>	177
II. Habla y literatura tradicional	321
Xus UGARTE BALLESTER,	

Les vins d'Anjou, les couilles de Lorraine et les Gentilz, qui furent buveurs eternelz: <i>ambivalencia y sociocentrismo lúdico</i> <i>en los blasones populares del Pantagruel de Rabelais,</i> <i>con muestras de versiones en varias lenguas.....</i>	323
Claude CARCENAC, <i>Femme et mariage à travers les sagesses de</i> <i>l'Égypte antique et les proverbes</i> <i>espagnols et français</i>	335
Cécile IGLESIAS, <i>La imagen del pícaro en los pliegos de cordel</i>	349
Ángel IGLESIAS OVEJERO, <i>Apotegmática:</i> <i>hechos y dichos tradicionales de El Rebollar</i>	381
José Benito MATEOS PASCUAL, <i>Riofrío, motes de las dos orillas</i>	411
Pilar TORRES VÁZQUEZ, <i>Son de antes</i>	423
Saturnino RODRÍGUEZ MIGUEL, <i>Intelectus apretatus.....</i>	429
III. Folclore y otras formas de cultura tradicional 441	
José Luis PUERTO, <i>La perspectiva celeste: leyenda, poesía y creencia</i> <i>en torno al cielo en el sur salmantino</i>	443
Pedro Javier CRUZ SÁNCHEZ, <i>Surmegir al santo. Las rogativas "pro pluvia" a San Ginés</i> <i>en Robleda (Salamanca)</i>	459
Lidio Jesús FERNÁNDEZ RODRÍGUEZ, <i>Manuel Llano y la mitología cántabra</i>	479
Rosa LORENZO LÓPEZ, <i>La mujer en la cultura tradicional.</i> <i>El drama de la madre soltera</i>	501
Carlos GARCÍA MEDINA, <i>Los trabajos y oficios desaparecidos</i> <i>de la tierra de Ciudad Rodrigo</i>	519
Gabriel CALVO, <i>Dendrolatrías familiares</i>	531
Ángel IGLESIAS OVEJERO, <i>El folclore fue historia vivida: Que rompan la Jesa Arriba.....</i>	547

UN EXEMPLE DE PROGRAMME REGIONAL : CRITERR

Rapport du Laboratoire Ligérien de Linguistique sur les actions inscrites dans le programme CRITERR – 2008 –

1. Observatoire du français et des contacts de langue en Région Centre

Ce programme, bénéficiant d'un soutien significatif de la part du Centre de ressources, a connu un développement très important au cours de l'année 2008, mobilisant une part essentielle des financements disponibles, soit :

- en matériel : achat d'équipement d'enregistrement et logiciels de traitement consommant les 9.000 € de crédits portés au crédit de cette action ;
- en personnel par la rémunération d'étudiants en vacation, chargés des opérations de saisie, transcription et relecture.

Pour mémoire, l'objectif est de constituer Orléans comme centre de référence pour la constitution d'un grand corpus de référence de français parlé (10.000.000 de mots) dont les applications concernent à parts égales la didactique (français langue maternelle / français langue seconde) et le Traitement Automatique des Langues (reconnaissance de la parole, fouille de données orales, indexation automatique...).

Le principal résultat du travail conduit cette année par le laboratoire concerne la constitution d'une base de données qui est d'ores et déjà la principale ressource en langue française parlée et qu'il convient d'incrémenter pour la porter au niveau requis par les exploitations visées. Afin d'assurer une accessibilité optimale aux données tout en réduisant les coûts de traitement, le LLL s'est rapproché du département de linguistique de l'Université de Louvain pour le formatage de données et du Laboratoire d'Informatique Mathématique et des Sciences pour l'Ingénieur (LIMSI - Paris XI/CNRS) pour mettre au point des techniques d'automatisation des procédures.

La phase technique de préparation des deux cents heures d'enregistrement est terminée, mais la nécessité d'une anonymisation interdit la mise en ligne qui sera l'aboutissement d'une démarche de partage des données sur le modèle de ce qui a pu être réalisé par l'ATILF à Nancy pour l'écrit. Pour le moment, et avec un accès réservé, les documents sont déposés à l'Université d'Orléans, dans la base ESLO (Enquête Socio-Linguistique à Orléans). Ils peuvent être présentés à des fins de démonstration et rendus accessibles pour des consultations contrôlées afin de respecter les exigences juridiques inhérentes à ce type d'enquête.

Cette base de données comprend en réalité trois enquêtes corrélées :

- une restitution archivistique d'une enquête conduite à Orléans à la fin des années 60,
- un renouvellement de cette enquête à des fins comparatives en cours de constitution (dynamique des langues),
- une étude des langues parlées à Orléans.

Chaque enregistrement est documenté, transcrit sous Transcriber et correspond aux prescriptions Dublin Core et OLAC. Il est associé à une documentation papier et à une mémoire des opérations conduites pour parvenir à l'élaboration de la base.

En termes de collaboration, la recherche associée, en France :

- la Bibliothèque nationale de France, Département des Documents Audiovisuels et Sonores (ex Phonothèque Nationale),

- l'Ecole des Hautes Etudes en Sciences Sociales (Centre de Linguistique Théorique),
- l'Institut de Linguistique Française qui fédère dix équipes CNRS travaillant sur la langue française et le LLL,
- l'Observatoire des Pratiques Linguistiques de la Délégation Générale à la Langue Française et aux Langues de France auprès du Ministère de la Culture et de la Communication,
- le LIMSI et
- le Laboratoire d'Informatique de l'Université de Tours (une collaboration qui doit être étendue au LIFO dès cette année).

A l'étranger, participent au projet des chercheurs des universités de Bonn (Allemagne), Louvain (Belgique), Exeter (Grande-Bretagne) et Limerick (Irlande).

Dans l'étude des langues parlées à Orléans, des études spécifiques ont été entreprises concernant les langues arabes, tamaghit, peul, anglaise, kriol... Une recension générale est en cours afin de mesurer d'une part la diversité des langues en usage dans l'agglomération et les influences réciproques entre le français et ces langues, dans les deux sens.

Des travaux ont été présentés au cours de cette année

- à Bâle (colloque sur les corpus romans, mai 2008, publication en cours, I. Eshkol),
- à Paris (Congrès Mondial de Linguistique Française, juin 2008, publication de G. Bergounioux consultable sur le site du colloque),
- à Lausanne (formations doctorales en linguistique de la Suisse romande, septembre 2008, publication de G. Bergounioux prévue en juin 2009 dans les *Cahiers de l'Institut de Linguistique*, Lausanne),
- à Oxford (*French Language Studies* Colloquium, septembre 2008, à paraître à OUP, contribution d'O. Baude), à Swansea (*Corpus Linguistics*, septembre 2008) et à
- à Oviedo (Congrès International de Linguistique, septembre 2008, à paraître dans les actes du colloque, contributions de L. Abouda, B. Hamma, F. Nemo et M. Petit).

D'autres présentations sont prévues dans les mois à venir.

2. Actions collaboratives (Guyane et Burkina Faso)

L'assistance apportée aux chercheurs en linguistique du Burkina Faso a conduit le laboratoire à accueillir deux doctorants issus du CNRST burkinabé dans l'objectif d'assurer, en plus d'un approfondissement des contenus disciplinaires, une formation aux techniques de collecte et de traitement des données. L'objectif est d'apporter une assistance scientifique et technique aux acteurs nationaux pour qu'ils puissent constituer progressivement une phonothèque nationale au standard international avec des copies de sauvegarde qui seraient hébergées par la BnF. Une part de l'accroissement de l'équipement informatique du laboratoire a été consacrée à ce projet.

Les travaux sur la Guyane sont en cours d'exploitation, sous la conduite d'un chercheur du laboratoire qui a été nommé à la direction du CNRS en Guyane. Une demande forte concerne la préparation des travaux qui seront publiés dans l'*Encyclopédie des langues et cultures de Guyane*, en partenariat entre les Presses Universitaires d'Orléans et le Comité des Travaux Historiques et Scientifiques (Ministère de l'Education Nationale). Sur les vingt ans à venir, c'est une soixantaine d'ouvrages qui seront publiés.

3. Comité pour la Simplification de la Langue Administrative en Région Centre

Décliné à partir de son homologue national, le COSLA, service du Premier Ministre destiné à l'orientation et à la prescription en matière de réforme de l'état, le COSLA-Centre a vocation à fonctionner comme un observatoire des pratiques administratives à même de vérifier à son niveau la mise en application des directives et comme un organe de recommandation. En 2008, un premier inventaire des besoins et des attentes a été dressé. La Région, par la voix de son Président, et le COSLA national ont manifesté leur intention de soutenir résolument cette initiative.

RESEAU FRANÇAIS DE PHONOLOGIE

Appel à communication pour le

Colloque du Réseau Français de Phonologie

In memoriam Nick Clements

Université d'Orléans

Jeudi 1, vendredi 2 et samedi 3 juillet 2010

Afin de prolonger la série des rencontres annuelles impulsées par le GDR « Phonologies », et dans le même esprit, le Réseau Français de Phonologie (RFP) lance un appel à communication pour un colloque qui se tiendra à l'Université d'Orléans du 1 au 3 juillet 2010, avec le soutien des deux fédérations de recherche du CNRS, Institut de Linguistique Française (ILF / FR 2393) et Typologie Universaux Linguistiques (TUL / FR 2559).

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'analyse du signal, la perception, l'acquisition, la diachronie, la dialectologie, les formalismes, l'épistémologie et toutes les questions qui s'inscrivent explicitement dans le domaine de la phonologie et de ses interfaces.

Format de soumission et sélection

Les propositions de communication, rédigées en français ou en anglais, seront présentées sous forme d'un résumé de deux pages maximum (format A4, police Times taille 12), incluant les éléments essentiels de bibliographie, et transmises à rfp@univ-orleans.fr en fichier Word, Latex ou PDF. Ce fichier ne fera aucune mention de l'auteur qui indiquera ses nom, civilité, adresse, institution et le titre de sa communication sur un fichier séparé. Les propositions seront transférées pour évaluation à deux rapporteurs et la sélection finale fera l'objet d'une discussion collective.

Calendrier

Appel à contributions: 1^{er} décembre 2009

Envoi des contributions avant le 1^{er} mars 2010

Retour des avis aux auteurs : 15 avril 2010

Colloque : du jeudi 1^{er} au samedi 3 juillet 2010, *in memoriam* Nick Clements.

(Les prochains colloques du RFP se tiendront à Tours en juin 2011 et à Paris en juin 2012)

Organisation

Les communications seront de 30 minutes, réductibles à 20 si le conférencier en fait la demande au moment de la soumission, à quoi s'ajouteront dans tous les cas dix minutes de discussion. Pas de session posters prévue. Deux conférenciers invités ouvriront les journées et une fin d'après-midi sera réservée à une assemblée générale du RFP.

Des sessions thématiques peuvent être proposées. L'organisateur est invité à donner un titre général avec un argumentaire de quelques lignes et la liste des collègues qui participeraient à la session avec un bref résumé de la thématique que chacun développerait. L'évaluation se fera lors de la réunion finale du comité de lecture sur proposition de deux rapporteurs et sans anonymat.

Le colloque ne prend en charge aucun frais, le coût de l'inscription couvrant les dépenses courantes. Aucune publication n'est prévue mais la question sera rediscutée au moment du colloque.

Comité de lecture

J.-P. Angoujard, S. Bendjaballah, G. Bergounioux, J. Brandao de Carvalho, Ph. Carr, M. D'Imperio, C. Dugua, J. Durand, S. Ferré, J.-M. Fournier, S. Herment, M.-P. d'Imperio, M. Lahrouchi, B. Laks, J.-L. Léonard, J. Lowenstamm, N. Nguyen, R. Noske, C. Patin, A. Rialland, R. Ridouane, M. Russo, P. Sauzet, T. Scheer, Ph. Ségéral, A. Tifrit, N. Vallée, S. Wauquier.

Comité d'organisation (Orléans & Tours)

V. Abasq, G. Bergounioux, C. Dugua, S. Ferré, J.-M. Fournier, N. Lampitelli, X. Luo, M. Martin

COLLOQUE RFP
Orléans 2010

Jeudi 1er juillet

10:00 - Allocutions de bienvenue

10:20 - Hommage à Nick Clements

10:40 - A. Rialland & M. Embanga Aborobongui : Le système intonatif de l'embosi

11:20 - T. Scheer : What the initial CV is initial of

13:30 - J. Pater : Systemic Simplicity

14:10 - S. Wauquier : Erreurs de liaison en acquisition du français L1 et L2

14:50 - C. Patin : De l'intérêt des domaines pour l'analyse tonale: tons et intonation en shingazidja

15:50 - N. Ballier & Ph. Martin : Corrélats prosodiques et acoustiques de la syllabification

16:30 - M. L. Mazzola : Aspects segmentaux du français régis par la prosodie

17:30 - R. Noske : Contrast Preservation Theory, Glottalic Theory and the Germanic Sound Laws

18:10 - M. Lahrouchi & Ph. Ségéral : Le gabarit en tagnawt

Vendredi 2 juillet

9:00 - E. Dresher : Conférence plénière: Feature Hierarchies Halle (1959) to Clements (2009)

10:00 - J. Bucci : La réduction vocalique en Coratin

11:00 - B. Laks : La phonotactique saussurienne : loi de la valeur relationnelle et positionnelle

11:40 - E. Caratini : "Divorce" in the history of German

11:40 - A. Fournet : La fusion de phonèmes en proto-arabe

13:40 - S. Ben Si Saïd : L'interaction entre la structure et la mélodie

13:40 - N. Arbaoui : Les multiples secrets de la forme II de l'arabe classique

14:20 - O. Piot : The prosody of complex intonation phrases in French

14:20 - N. Yamaguchi : La fréquence des traits peut-elle prédire l'ordre d'acquisition des consonnes?

15:00 - K. Hammane : Miroir, miroir, dis-moi à quoi ressemble une consonne syllabique...

15:00 - N. Lampitelli : Du caractère abstrait de la composante phonologique

16:00 - J.-M. Fournier : Accentuation lexicale et poids syllabique en anglais

16:40 - S. Bendjaballah & Ph. Ségéral : Le système vocalique et l'accent en mehri.

17:20 - O. Rizzolo : Une énième ballade sur les chemins du verlan

18:20 Assemblée générale du RFP

Samedi 3 juillet

9:00 - J.-L. Léonard : Classes naturelles, traits, complexité et simplicité segmentale en tu'un savi (mixtec)

9:40 - A. Delplanque : Les consonnes finales du dagara: propriétés des Eléments Phonologiques

10:20 - R. Välimaa-Blum : Why it is so difficult to violate vowel harmony in Finnish: A cognitive view

11:20 - J. Brandão de Carvalho : Je ll'ai vu : the case for abstract candidates

12:00 - J. Berns : Bruits fricatifs en français parisien : le dévoisement vocalique et l'affrication plosive

14:00 - J. Lowenstamm : Cycles et phases

14:40 - J.-M. Beltzung : Dissimilation consonantique par effacement dans certaines langues bantu

15:20 - N. Faust : Où est le [u]? Sur les conséquences théoriques de la décomposition de morphèmes

16:00 Conclusion

1. Présentation du projet

Identité du coordonnateur du projet :

Nom – Prénom : Caron Philippe et Hanote Sylvie

Corps et grade : PR et MCF

Laboratoire d'affiliation du coordinateur : FORELL A EA3816

Courriel : philippe.caron@univ-poitiers.fr ; sylvie.hanote@univ-poitiers.fr

Téléphone : 05.49.45.46.51 (Marie-Christine Mérine, secrétariat FORELL)

Nom du projet : « Parole »

Mots-clés (5 maximum) : Phonétique, prosodie, analyse acoustique, synthèse de la parole

Type de projet

- Lié à une action existante du présent contrat quadriennal (CPER, ACI, ANR,...)
- Déjà élaboré (en cours de soumission auprès d'un organisme financeur)
Précisez lequel :
- En cours d'élaboration

Concernant le volet « phonétique anglaise et analyse acoustique », une ACI Université de Poitiers est en cours (2009-2010). Cette ACI réunit deux laboratoires de l'université de Poitiers : FORELL A (EA3816, MSHS Poitiers) et le laboratoire XLIM-SIC (UMR CNRS 6172), en association avec la fédération FELINCO (Poitiers, Limoges, Orléans, Tours). Cette ACI s'appuie sur le plateau technique « Analyse du signal sonore et traitement de la parole » de la MSHS. (Budget alloué : 12 000 euros.)

Objet d'étude et approche scientifique

Le projet « Parole » met en oeuvre des connaissances (phonétiques, prosodiques) et des outils technologiques au service de deux programmes complémentaires :

- l'analyse acoustique par ondelettes (avec l'anglais oral contemporain comme principale application) permet de pénétrer plus avant dans la réalisation orale et en élucide plus finement les caractéristiques ;
- la restitution de la diction haute du français classique vers 1700 (synthèse de parole, logiciel Kali) est un cas d'archéologie phonético-prosodique intéressant puisqu'il s'agit de faire revivre un objet phonétique et esthétique perdu.

L'objet d'étude commun est la parole, le projet est l'élaboration de corpus oraux et le recours à des outils de traitement du signal pour analyser ou reconstruire la parole. Les deux axes sont donc complémentaires et s'enrichissent mutuellement.

Etat de l'art et référence clefs :

Organisation de 2 journées d'études à la MSHS de Poitiers (Octobre 2008 et octobre 2009).

- o Octobre 2008 : journée « Parole » co-organisée par P. Caron et S. Hanote, où sont intervenus des spécialistes de la mise en place de corpus oraux (présentation du corpus *Eslos*, LLL Orléans), des spécialistes de l'analyse de la parole et de la synthèse de parole.

- Octobre 2009 : journée « Intensité » co-organisée par S. Hanote et S. Moore, où sont intervenus des spécialistes de la phonétique acoustique (chercheur invité : P. Martin Université Paris-Diderot EA3967), avec une table ronde finale sur la question « L'intensité, est-elle une mesure pertinente en analyse de la parole ? »

La publication des travaux de ces deux journées d'études est prévue dans la revue électronique *CORELA – Cognition, Représentation, Langage*. <<http://revue-corela.org>>. La première paraîtra en juin 2010, la seconde à l'automne 2010.

Problématique et méthodologie :

- Le programme 'restitution de la diction haute du français vers 1700' s'impose actuellement car il répond à un besoin chez les artistes, musiciens et acteurs, qui sont demandeurs de connaissances avérées. D'autre part il est le seul de son type actuellement et il permettra d'éviter que se développent des idées originales mais scientifiquement mal assises, dérivées de l'œuvre pionnière mais entachée d'erreurs d'Eugène Green (*La Parole baroque*, Desclée de Brouwer 2002).

La méthodologie dans ce programme consiste à collationner dans les manuels (grammaires, dictionnaires généraux et dictionnaires de rimes, manuels de français oral etc.) et les remarques du temps un maximum de données et d'indices relatifs à la phonétique et à la prosodie de cet habitus, de les comparer, d'essayer de les comprendre et de les interpréter. Puis l'enregistrement par un acteur d'un corpus oral simulé va permettre d'extraire à la fois les segments minimaux de parole (les diphtongues), segments qui seront ensuite réagencés librement par la synthèse de parole à la demande d'un usager qui demandera l'oralisation-type d'un texte de son choix. Puis ce texte, vers par vers, sera analysé et sa structure métrico-syntaxique se verra dotée, sur la base de ses propriétés, d'un patron prosodique le plus proche possible de sa structure.

- En ce qui concerne le versant « analyse acoustique et phonétique anglaise », le travail commencé dans le cadre de l'ACI (2009-10) a d'ores et déjà permis :

- le début de la mise en place d'un corpus d'anglais oral authentique (enregistrements de locuteurs d'une part et enregistrements radiophoniques d'autre part),
- le développement du logiciel de traitement de la parole *Speech Wave Analyser* dans le cadre de la coopération avec le laboratoire XLIM-SIC.

Il s'agira de poursuivre le travail commencé sur ces deux versants et de valider le logiciel sur des problématiques de phonétique anglaise (la préfixation, la liaison, la prononciation des diphtongues...). L'objectif est de faire ressortir des signatures acoustiques en accord avec les phénomènes phonétiques / phonologiques.

Programme (organisation temporelle des étapes) :

Analyse acoustique et phonétique anglaise :

- a) Extension du corpus d'anglais oral et mise en place d'un mode d'interrogation de ce type de corpus. Pour affiner les contours scientifiques et techniques de cette étape, des contacts sont pris avec des laboratoires qui travaillent déjà sur la mise en place d'un corpus oral (corpus Aix-MARSEC Machine Readable Spoken English à Toulouse Le Mirail, Corpus PFC (Phonologie du Français Contemporain) et maintenant PAC (Phonologie de l'Anglais Contemporain : Jacques Durand et Bernard Laks, entre autres, à Toulouse également). De plus, pour la mise en place d'un mode d'interrogation efficace de ce type de corpus, l'intervention d'un ingénieur d'étude est nécessaire. A ce titre, un dialogue est établi avec Julien Eychenne, qui a travaillé sur la plateforme PFC pour qu'il nous fasse bénéficier de son expérience sur le corpus PFC. (24 mois)
- b) Test du logiciel *Speech Wave Analyser* sur des phénomènes de phonétique anglais pertinents. Dans sa thèse (2009-) « Préfixation et phonologie de l'anglais : analyse lexicographique, phonétique et acoustique » (direction : Jean-Louis Duchet, co-direction

Sylvie Hanote), Nicolas Videau va travailler sur la prononciation des préfixes en anglais oral contemporain. Actuellement, le logiciel a été développé par Philippe Carré (responsable équipe de recherche ICONES, laboratoire XLIM-SIC), puis par Raphaël Soulard (Doctorant XLIM-SIC) dans le cadre de l'ACI. Maintenant que les bases du logiciel sont posées, et en parallèle du travail de recherche fondamentale sur les phénomènes de phonétique anglaise, une partie technologique doit être assurée par un ingénieur pour faire évoluer le logiciel et assurer sa diffusion au sein de la communauté. En effet, nous souhaitons, à moyen terme, proposer une solution logicielle aux laboratoires travaillant sur des sujets connexes. (24 mois)

Ce projet d'analyse acoustique et phonétique anglaise s'articule avec un autre projet avec le Laboratoire Ligérien de Linguistique (LLL Tours) sur la numérisation de données phonétiques (J.M. Fournier (LLL) et J.L. Duchet (FORELL-A)).

Restitution de la diction haute du français

2010 : journée Parole III, destinée à agréger autour des laboratoires cocontractants (FORELL Poitiers, CRISCO de Caen et CESR de Tours) des partenaires étrangers spécialistes des manuels de français langue étrangère au XVIIe siècle et des experts (Yves-Charles Morin, Université de Montréal et Benoit de Cornulier, Université de Nantes). Concertation en vue de la rédaction d'un livre collectif qui devrait faire autorité.

2011 : enregistrement du corpus de base par Olivier Bettens (Lausanne) et/ou Pierre-Alain Clerc (Genève). Prélèvement des diphones et ajout de logatomes pour les combinaisons absentes du corpus

2012 : création d'un glossaire phonétique des mots spécifiques du registre grave (tragédie dramatique et lyrique)

2013 Premiers tests du logiciel de synthèse de parole. Correction des erreurs phonétiques et prosodiques.

2014 Création de l'environnement scientifique de l'outil en ligne (c'est-à-dire les pages fixes qui constitueront la mise en perspective des résultats acquis par le programme et qui s'incarnent dans la synthèse de parole).

2015 Installation de la base en ligne, en libre accès.

2. Partenaires, ressources humaines et matérielles

Partenaires du projet :

- Scientifiques (labos – MSH – universités...)
 - FORELL-A EA3816 (MSHS Poitiers)
 - CRISCO (MSHS de Caen)
 - X-LIM SIC UMR CNRS 6172, Université de Poitiers / Limoges
 - CeReS (Université Limoges)
 - LLL (future intégration dans la MSHS Tours)

Ressources humaines :

Nom et prénom	Statut (PU, MCF, CR, DR, ...)	Laboratoire	Organisme d'appartenance (CNRS, Université de Poitiers, ...)	Taux de participation (ETP) ¹
Caron Philippe	PR	FORELL -A	Université Poitiers	0,6
Duchet Jean-Louis	PR	FORELL -A	Université Poitiers	0,2
Col Gilles	PR	LLL	Université Tours	0,3
Hanote Sylvie	MCF	FORELL -A	Université Poitiers	0,6
Moore Susan	MCF	CeReS	Université Limoges	0,3

Videau Nicolas	Doctorant	FORELL- A	Université Poitiers	0,6
Morel Michel	Ingénieur	CRISCO	Université Caen	0,4
Laurenti Jean-Noël	MCF	CESR	Université Tours	0,3
Delente Eliane	MCF	CRISCO	Université Caen	

En outre, des partenaires étrangers participent au projet et viendront à la journée PAROLE III, destinée à rendre leur participation plus organique :

- Yves-Charles Morin (Université de Montréal)
- Olivier Bettens (chercheur indépendant, spécialiste de musique vocale et de phonétique historique (Lausanne))
- Elmar Schafroth (Université de Düsseldorf)
- Sergueï Vlassov (Université de Saint Petersburg)

Ressources matérielles :

Utilisation des plateaux techniques MSHS :

Oui Non

Si oui lequel :

Ressources numériques

Analyse du comportement humain en temps réel

Plateau technique « Parole » pour l'enregistrement de locuteurs et le traitement des données.

3. Valorisation

Productions envisagées (revues, ouvrages, base de données...)

Corpus oral anglais contemporain en discours avec mode d'interrogation automatique.

Logiciel de traitement de la parole 'Speech Wave Analyser'.

Publications dans revues.

Colloques et manifestations

Journée PAROLE III (automne 2010), réservée au volet « restitution de la parole ».

4. Nature du soutien demandé à la MSHS

Projet porté par la MSHS (soutien financier, technique et administratif)