

**Laboratoire**  
**Modèles, Dynamiques, Corpus**

**Unité mixte de recherche 7114**

**Responsable : Bernard Laks**

**Départements : SHS et MPPU**

**CNRS**  
**Université Paris 10**  
**Université Paris 5**

**Rapport quadriennal 2005-2008**  
**Projet 2009-2012**

**Volume 3/3**

**15 octobre 2007**

# Table des matières

---

<b>II. DOSSIER SCIENTIFIQUE .....</b>	<b>3</b>
3. Projet scientifique pour la période 2009-2012.....	3
3.1. Déclaration de politique scientifique pour la période 2009-2012.....	3
3.2. Structure proposée du laboratoire - Organigramme.....	10
3.2.1. Organigramme administratif .....	10
3.2.2. Organigramme scientifique .....	11
3.2.3. Récapitulatif quantitatif pour le projet 2009-2012 .....	12
3.3. Projet scientifique - période 2009-2012.....	13
3.3.1. Composante <i>Modèles</i> .....	14
3.3.1.1. Equipe <i>Grammaires</i> .....	15
3.3.1.1.1. <i>Lexique, morphologie, sémantique</i> .....	18
3.3.1.1.2. <i>Modèles formels et traitement automatique</i> .....	22
3.3.1.1.3. <i>Phonologie de corpus et phonologies dynamiques</i> .....	25
3.3.1.1.4. <i>Syntaxe et sémantique grammaticale descriptives</i> .....	29
3.3.1.2. Equipe <i>Activité langagière</i> .....	33
3.3.1.2.1. <i>Unité(s) du texte</i> .....	34
3.3.1.2.2. <i>Genres</i> .....	38
3.3.1.2.3. <i>Expression</i> .....	43
3.3.2. Composante <i>Dynamiques</i> .....	49
3.3.2.1. Equipe <i>Dynamiques des acquisitions</i> .....	50
3.3.2.1.1. <i>Acquisition initiale du langage</i> .....	51
3.3.2.1.2. <i>Acquisition des langues secondes ou étrangères et du plurilinguisme</i> .....	56
3.3.2.1.3. <i>Pathologies de la communication et du langage et remédiations</i> .....	59
3.3.2.2. Equipe <i>Dynamiques des usages</i> .....	65
3.3.2.2.1. <i>Variétés des styles et des dynamiques de la production orale</i> .....	66
3.3.2.2.2. <i>Variétés des styles et des dynamiques de la production écrite</i> .....	69
3.3.3. Opérations inter-équipes.....	73
3.3.3.1. <i>Corpus de langage oral : formats, conversion, visualisation et interrogation</i> .....	73
3.3.3.2. <i>Variétés des genres discursifs scolaires et universitaires</i> .....	75
3.3.3.3. <i>Dynamiques acquisitionnelles et sociolinguistiques des situations de la francophonie périphérique</i> .....	77
3.3.3.4. <i>Interactions langagières</i> .....	80

## II. DOSSIER SCIENTIFIQUE

---

### 3. PROJET SCIENTIFIQUE POUR LA PERIODE 2009-2012

(comprenant l'organigramme de l'unité prévu pour 2009)

#### 3.1. DECLARATION DE POLITIQUE SCIENTIFIQUE POUR LA PERIODE 2009-2012

Le laboratoire Modèles, Dynamiques, Corpus, UMR 7114 est placé sous la triple tutelle du CNRS, de Université de Paris 10 (tutelle principale) et de l'Université de Paris 5 (tutelle secondaire). Créé en janvier 2001 il a été renouvelé en janvier 2005 et sollicite un nouveau renouvellement pour la période 2008-2012.

Le laboratoire inscrit son activité dans la continuité thématique et méthodologique de la recherche linguistique menée à l'Université Paris 10 depuis plusieurs décennies. Centrée sur la linguistique française, sans pourtant exclure *a priori* l'étude d'autres langues, cette unité thématique est marquée par un souci empirique qui adosse les analyses linguistiques à une base descriptive solide prenant en compte la réalité des usages, qu'ils soient conversationnels, discursifs ou textuels. C'est donc sur la base de corpus de tous ordres et de toutes dimensions que sont proposées des analyses et des modélisations linguistiques s'inscrivant dans des cadres théoriques modernes assez divers. L'unité thématique et méthodologique du laboratoire se double en effet d'une grande ouverture théorique et d'une polyphonie revendiquée pour ce qui concerne la modélisation. La spécificité des travaux du laboratoire *MoDyCo* consiste ainsi à intervenir dans les débats théoriques les plus fondamentaux à partir d'un ancrage solide en linguistique française et d'un souci empirique et descriptif constant.

Dans la dernière période, le laboratoire a ainsi développé, seul ou en partenariat, de nombreux corpus linguistiques. On peut citer le « corpus PFC » les corpus construits dans le cadre du projet PFC (resp. B. Laks), le corpus « Hommes politiques » (resp. B. Laks), les corpus mis à disposition par Jean Dubois et Françoise Dubois-Charlier (resp. D. Leeman et D. Le Pesant), le corpus « Duras » (resp. D. Malrieu).

Construisant ces corpus, le plus souvent avec des partenaires externes, le laboratoire s'est attaché à mettre en place une politique de visibilité et d'accessibilité sur internet particulièrement active. Il a développé plusieurs site spécifiques et mis en place une politique très positive de production d'outils spécialisés (fouille, indexation, recherche automatique etc.).

Au cours du dernier contrat, le potentiel scientifique du laboratoire s'est trouvé considérablement renforcé par l'intégration de deux composantes nouvelles. En 2005, l'intégration du CELITH (tutelle secondaire avec l'EHESS) a permis de combler les manques en sémantique et pragmatique et a considérablement renforcé les recherches sur les nouveaux médias et les nouveaux modes d'écriture. Sur cette même thématique de recherche, le soutien

de l'université Paris 10 au laboratoire a été marqué par l'attribution de la mise au concours d'un poste d'enseignant-chercheur en sémantique au titre des recrutements exceptionnels de 2006 destinés à la recherche et au renforcement des laboratoires.

En 2006, l'intégration du LEAPLE (tutelle secondaire avec Paris 5) a permis de donner une nouvelle ampleur aux recherches en acquisition, normale et pathologique, en apprentissage précoce en milieu familial ou scolaire et en apprentissage de langue seconde.

Le renforcement du laboratoire dans les domaines de la sémantique, de la pragmatique et de la psycholinguistique a conduit à une réorganisation de la structure interne du laboratoire en deux composantes scientifiques plus clairement identifiées, constituées chacune de deux équipes :

1) Modèles : équipe « Grammaires » et équipe « Activité langagière » ;

2) Dynamiques : équipe « Dynamiques des acquisitions » et équipe « Dynamiques des usages ».

Les corpus, bases de données, plateformes et outils sont conçus comme transverses à ces deux composantes majeures.

Le projet soumis pour la période 2009-2012 s'appuie sur cette réorganisation interne et identifie aussi bien les spécificités de chacune des composantes et des équipes qui les constituent que les synergies nombreuses existant entre les actions menées.

[...]

### **Thématiques de recherche**

Les thématiques de recherche conduites au sein du laboratoire se déclinent selon trois grandes options :

1) La perspective *modélisatrice* concerne au premier chef les recherches phonologiques, morphologiques, syntaxiques et sémantiques, ainsi que les travaux de formalisation en traitement automatique des langues. C'est l'objet des recherches menées par l'équipe « Grammaires ». L'équipe « Activité langagière » est concernée quant à elle par les thématiques pragmatiques, discursives, textuelles et formes symboliques.

La perspective *modélisatrice* se donne comme programme de recherche la description explicite et formalisée des langues, qui se réalise par la définition et l'exploration de différents plans organisationnels du langage (à tout le moins : phonologie, morphologie, syntaxe, sémantique) et par la découverte des relations de dépendance et d'autonomie entre ces plans. En effet un point de convergence pour la plupart des chercheurs peut se résumer en la formule « le langage n'est pas un » ou « le langage est structurellement hétérogène ».

2) La perspective *dynamiques* structure deux équipes : l'équipe « Dynamiques des acquisitions » et l'équipe « Dynamiques des usages ». La première est concernée par les thématiques d'acquisition de langue maternelle ou seconde, d'acquisition normale et pathologique, en milieu familial ou scolaire ainsi qu'en didactique. La seconde est concernée par les thématiques sociolinguistiques et sociostylistiques, par les problèmes liés à la variation ainsi que par les spécificités de l'oral et les rapports à la norme.

La prise en compte des usages et des dynamiques contribue à l'évaluation et à l'évolution des modèles linguistiques. Des études récentes sur l'acquisition des langues illustrent la

compétition entre fonctions phrastiques, sémantiques et discursives. Les recherches sur l'acquisition et sur les usages en situation obligent aussi à tenir compte de la variabilité des systèmes linguistiques à travers le temps, l'espace et les caractéristiques sociales des locuteurs, et à rendre compte de la façon dont chaque locuteur fait sienne une langue pour ses visées communicatives concrètes, régies par des contraintes cognitives, situationnelles et interactionnelles. Elles visent à décrire comment une langue est couramment utilisée par le locuteur, comment elle peut être construite par l'enfant devenant locuteur, ou par un sujet acquérant un nouveau système linguistique. Elles amènent enfin à examiner comment les comportements linguistiques et l'appropriation de la langue sont en partie sous-tendus par les représentations que les locuteurs se font de la langue, y compris celles induites ou construites par l'institution éducative.

3) Transverse, la perspective *corpus*, tout en étant plus spécifique des recherches en linguistique textuelle, diachronique ou discursive, et en linguistique de l'écrit, concerne, en tant que méthodologie d'appui, les autres perspectives de recherche.

L'étude des corpus en situation montre que le lexique, la morphologie, la syntaxe, la manière dont se posent les problèmes sémantiques de l'ambiguïté et de l'implicite, tout cela varie aussi avec les genres et les discours. Pour parvenir à des traitements automatiques spécifiques et efficaces de corpus, il convient de spécifier les fonctionnements propres aux différents discours et genres textuels pour adapter les stratégies. Par ailleurs, la typologie textuelle peut apporter de nouvelles données à la morphologie et à la syntaxe. Bref, l'étude des normes d'usage complète utilement celle des règles, et permet de préciser leurs conditions d'application.

L'analyse de grands corpus intervient alors tant dans la formulation des hypothèses que dans la recherche de contre-exemples ou de variations. On découvre à l'intérieur d'une même langue des normes de portée et de niveau différents (de discours, de genres, de styles). Ainsi, le statut même de l'empirique est-il en train de changer. Ainsi, de par les thématiques qu'il travaille, le laboratoire *MoDyCo* est à même de prendre part à ce débat proprement théorique.

La constitution de corpus informatisés de données orales, associés à des techniques nouvelles d'étiquetage et de pré-analyse autorisant des traitements à grande échelle, permet ainsi des avancées dans la validation des hypothèses sur le rôle respectif de différents niveaux d'analyse et niveaux d'organisation de la langue, et engage en outre à modifier le rapport entre théorie et pratique.

Ces trois dimensions de recherche ont structuré et continuent de structurer les opérations en cours au sein du laboratoire. Les différentes opérations, à travers desquelles un grand nombre de sous-disciplines des sciences du langage se trouvent convoquées, montrent comment se trouvent appliqués à des objets linguistiques les plus divers les mêmes soucis d'ordre méthodologique, descriptif et modélisateur : la diversité des objets linguistiques traités marque la richesse des recherches menées au sein du laboratoire, l'attention portée aux dynamiques langagières et acquisitionnelles, le souci descriptif et empirique dans la construction de corpus et la préoccupation de formalisation et de modélisation des hypothèses et des résultats soulignent les synergies et signent l'unité du laboratoire *MoDyCo*.

[...]

### **Recherche contractualisée**

Le laboratoire *MoDyCo* s'est fixé pour le prochain contrat quadriennal un certain nombre d'axes majeurs dans lesquels il compte poursuivre une recherche d'excellence. Il compte notamment accentuer sa politique de recherche en partenariat national et international.

Depuis 2005, date de création de l'ANR, le laboratoire a été particulièrement attentif aux axes privilégiés par l'agence et a été très réactif à ses appels d'offres.

Les nombreux projets pilotés par *MoDyCo*, ou dans lesquels le laboratoire est partenaire, ont connu une évaluation très positive, modifiant très sensiblement l'équilibre scientifique et financier du laboratoire.

En 2007, notre laboratoire est pilote ou partenaires de 8 projets ANR. Parallèlement, les membres du laboratoire sont engagés dans des projets financés par l'ILF, la Fondation de France, le Ministère de la Culture, la DGLFLF, le Conseil scientifique de l'université Paris 10, etc. (cf. § 1.3.2., partie Bilan, « Les ressources propres » pour la liste des projets financés à hauteur de plus de 3 000 euros).

La réponse aux appels d'offres, nationaux et européens est ainsi devenue un axe privilégié de la politique scientifique du laboratoire. Une information systématique des enseignants-chercheurs, une aide à la réflexion stratégique, un soutien logistique et la mise en place d'un véritable incubateur de projets en sont les signes effectifs. La direction du laboratoire et le conseil de laboratoire ont ainsi mis en place une petite ligne budgétaire permettant de faire mûrir partenariats et projets. Cette politique volontariste sera développée dès que seront servis les préciputs correspondants aux projets en cours d'achèvement. Le développement et l'approfondissement de liens étroits avec d'autres établissements de recherche (Universités, Unités Propres du CNRS, Unités Mixtes du CNRS, INSERM, INRIA, CEA) s'inscrit dans ce cadre.

Le site du laboratoire (intranet et internet) permet de dynamiser cette politique de soutien à projets à laquelle participent activement les Ingénieurs et Techniciens ainsi que le secrétariat et la gestionnaire des contrats.

### **La politique de formation vis-à-vis des doctorants**

La politique de formation vis-à-vis des doctorants menée par le laboratoire *MoDyCo* s'inscrit dans une relation très étroite avec l'Ecole Doctorale 139.

Nous entendons la poursuivre et la développer sur la base des initiatives qui seront proposées.

Pour les thèses en cotutelle, nous entendons poursuivre notre politique d'incitation via les politiques mises en place par l'Ecole doctorale « Connaissance, langage, modélisation ».

Nous devons par contre faire un effort sur le plan de la durée de la thèse. Si trois ans paraît une durée difficile à tenir, nous souhaiterions que les thèses aient une durée de maximum 4 ans.

De même, nous devons aussi améliorer la répartition des thèses / directeurs de thèse, et ce sur deux plans :

- en favorisant la direction et la co-direction de thèses avec des chercheurs ;
- en répartissant mieux les directions de thèse entre les HDR. Nous souhaiterions qu'à l'avenir un directeur de thèse encadre au maximum 5 thèses.

Enfin, si depuis la réforme du LMD, mais aussi pour assurer ses cohérences thématiques, le laboratoire a opéré un recentrage sur les sujets de thèse des doctorants (ils doivent désormais s'inscrire dans les opérations menées par le laboratoire), cette politique n'a réellement débuté

qu'à l'automne 2006. Nous comptons la poursuivre et la développer avec le concours de l'Equipe de formation doctorale.

### **Les projets de collaboration internationale**

Depuis sa fondation en 2001, le laboratoire *MoDyCo* déploie une intense activité internationale. La direction du laboratoire est particulièrement sensible à cette dimension. Outre la continuation des nombreuses collaborations internationales déjà engagées (cf. partie « Bilan », § 1.8.), les projets du laboratoire s'inscrivent dans le cadre de l'extension de l'Ecole Doctorale internationale construite par l'Ecole doctorale 139 (extension à la Suède en 2008 à l'initiative de *MoDyCo*, avec des collaborations en linguistique du texte, extension à la Grande Bretagne projetée).

Dans le cadre de la politique contractuelle, on notera la prolongation du partenariat avec la SNS de Pise (Contrat UFI), avec l'Université de Chicago (Fond Franco-Américain pour la recherche), et avec l'université d'Ottawa.

Les contrats ANR encadrés par *MoDyCo* dans la prochaine période incluent tous une dimension européenne ou internationale (cf. partie « Bilan », § 1.3.2.).

Le projet Microbio, financé par STIC-AMSUD, (resp. Jean-Luc Minel) a permis de nouer de nouvelles collaborations avec 4 pays d'Amérique Latine.

### **Politique de diffusion et valorisation des connaissances**

#### **1) Les sites**

Nous entendons poursuivre notre politique de diffusion et de valorisation des connaissances. Nous souhaiterions que dans les années qui viennent chaque contrat ou chaque opération d'envergure développe son propre site.

Pour le site *MoDyCo*, dans la continuité de ce que nous avons commencé, nous entendons :

- continuer de développer les rubriques de l'intranet, de telle sorte que le site devienne pour chacun des membres du laboratoire un outil d'information et de diffusion.
- privilégier l'actualité ; la page d'accueil est conçue comme une page dans laquelle le nombre et l'agencement des informations peuvent varier en fonction de l'actualité du laboratoire, de la communauté et des institutions ;
- mettre à disposition de la communauté un ensemble de ressources. L'un des prochains développements concerne la consultation en ligne par la communauté des corpus de Jean Dubois et Françoise Dubois-Charlier (dans le cadre de l'opération « Syntaxe et sémantique descriptives du français » (resp. D. Leeman)), soit plus de 32 000 entrées.

#### **2) Le séminaire de *MoDyCo***

Le séminaire organisé par le laboratoire et l'école doctorale « Connaissance, langage, modélisation » sera animé par les équipes du laboratoire avec la volonté de constituer tout à la fois un lieu de diffusion des connaissances et de la diversité des méthodologies de la recherche en sciences du langage, et de brassage entre les différentes générations et statuts des membres du laboratoire. Il réunira régulièrement (9 séances par an) sur les thématiques d'excellence des différentes équipes du laboratoires les interventions centrées autour d'un axe de recherche bien délimité (a) un chercheur spécialiste bien identifié du domaine, et (b) des chercheurs ou enseignants-chercheurs et les doctorants relevant de l'opération responsable de la séance. En principe, chaque équipe devrait prendre la responsabilité de deux à trois séances

organisées de façon cohérente afin d'éviter l'atomisation de l'information et de permettre un véritable suivi des échanges entre intervenants et membre des autres équipes du laboratoire. Ceci devrait permettre à chacune des équipes du laboratoire de s'exprimer périodiquement tout en renforçant la cohésion et les échanges internes du laboratoire. Chaque année, une sélection des interventions donnera lieu à une publication dans les Cahiers de l'école doctorale, permettant ainsi au laboratoire de jouer pleinement rôle de formation par et pour la recherche.

### **3) Les CALaP**

Les CALaP continueront d'être publiés : ils permettent de valoriser le travail de l'une des équipes du laboratoire. Les coûts de fabrication et de diffusion étant très élevés, nous travaillons, comme l'ont fait beaucoup d'autres revues, à la transformer en revue électronique. Le site de *MoDyCo* sera le lieu d'accueil de la revue.

### **4) Le colloque jeunes chercheurs Coldoc**

Le laboratoire *MoDyCo* propose chaque année depuis 2004 un colloque ouvert aux doctorants et jeunes chercheurs, Coldoc. Cette manifestation a lieu à l'université Paris 10 et se déroule sur deux jours.

L'organisation est laissée à des doctorants et jeunes chercheurs du laboratoire. De la rédaction de l'appel à communication à la sélection du comité scientifique responsable de l'évaluation des propositions de communication, toutes les étapes de l'organisation sont prises en charge par le groupe de doctorants et jeunes chercheurs du laboratoire *MoDyCo* constitué en comité d'organisation. Jusqu'à présent, les thèmes suivants ont été abordés :

2004 – *La construction des observables en sciences du langage*

2005 – *Recueil de données en sciences du langage et constitution de corpus : données, méthode, outillage.*

2006 – *Intra-disciplinarité et extra-disciplinarité en sciences du langage*

2007 – *Le vocabulaire scientifique et technique en sciences du langage*

Chaque année, priorité a été donnée à un thème fédérateur autour duquel des doctorants de tous les champs des sciences du langage peuvent se retrouver.

Le site internet du laboratoire héberge les actes de Coldoc, ce qui permet aux doctorants et jeunes chercheurs participant d'accéder à une publication.

Le laboratoire *MoDyCo* soutient cette initiative en tant qu'elle constitue un élément important dans la formation des doctorants. D'une part, l'organisation permet aux doctorants et jeunes chercheurs impliqués de se former à l'organisation d'un tel événement (recherche de fonds, constitution d'un comité scientifique, etc.) ; d'autre part, Coldoc est un lieu de diffusion des travaux des doctorants et jeunes chercheurs ce qui permet ainsi une première expérience d'intervention publique et/ou publication (ce second aspect concerne l'ensemble de la communauté des doctorants et jeunes chercheurs puisque l'appel à communication est diffusé à l'échelle nationale).

Les différentes éditions de Coldoc ont toutes rencontré un certain succès, et ce constat encourage le renouvellement de l'opération. Chaque année, les soumissions de proposition de communication sont plus nombreuses et les prétendants à l'organisation ne manquent pas. Ceux-ci bénéficient d'un passage de relais : le comité d'organisation de l'année précédente



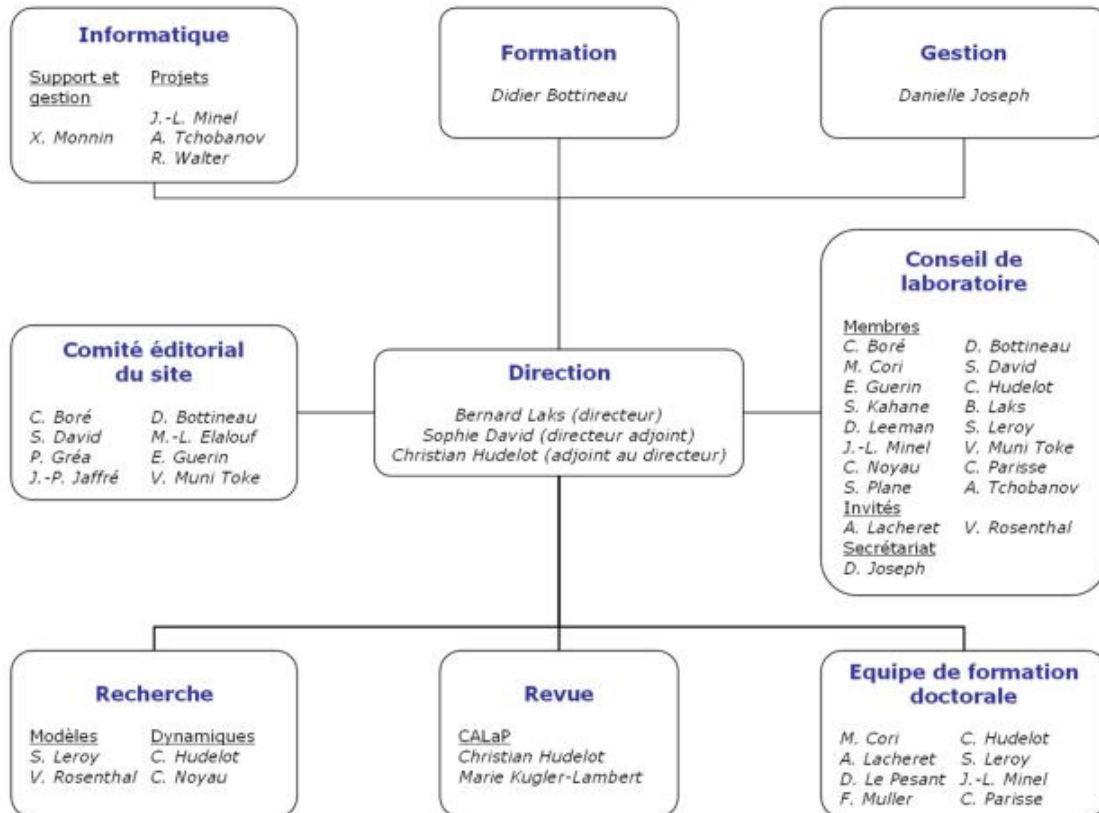
renseigne le nouveau sur les démarches à suivre. Ainsi, si les chercheurs ou enseignants chercheurs du laboratoire apportent leur soutien en encadrant les organisateurs, il n'en reste pas moins que ce colloque reste à la charge des doctorants et jeunes chercheurs.

Étant donné la place qu'attribue le laboratoire *MoDyCo* à la formation des doctorants, tout est mis en place pour que de prochaines éditions de Coldoc, enrichies des expériences passées, voient le jour.

[...]

## 3.2. STRUCTURE PROPOSEE DU LABORATOIRE - ORGANIGRAMME

### 3.2.1. ORGANIGRAMME ADMINISTRATIF



#### Activités en matière de gestion et d'informatique

##### A) Gestion

###### 1) Gestion financière de l'unité (D. Joseph)

- Gestion des contrats ANR
- Gestion des missions
- Gestion des commandes

###### 2) Gestion des personnels (D. Joseph)

##### B) Informatique

###### 1) Informatique de support et de gestion (X. Monnin)

- \* Administrateur du site du laboratoire.
- \* Gestion/maintenance du parc informatique : postes de travail, impressions, serveurs (Windows/Linux), stockage/sauvegarde.
- \* Assistance utilisateurs : systèmes (Windows, Linux, MacOS), logiciels de bureautique, logiciels de traitements audio/vidéo.
- \* Responsable des inventaires (matériels informatiques et petits matériels).
- \* Responsable d'achats : études, conseils avant vente et devis pour l'équipement informatique du laboratoire (ordinateurs, serveurs et périphériques), du matériel audio et vidéo.

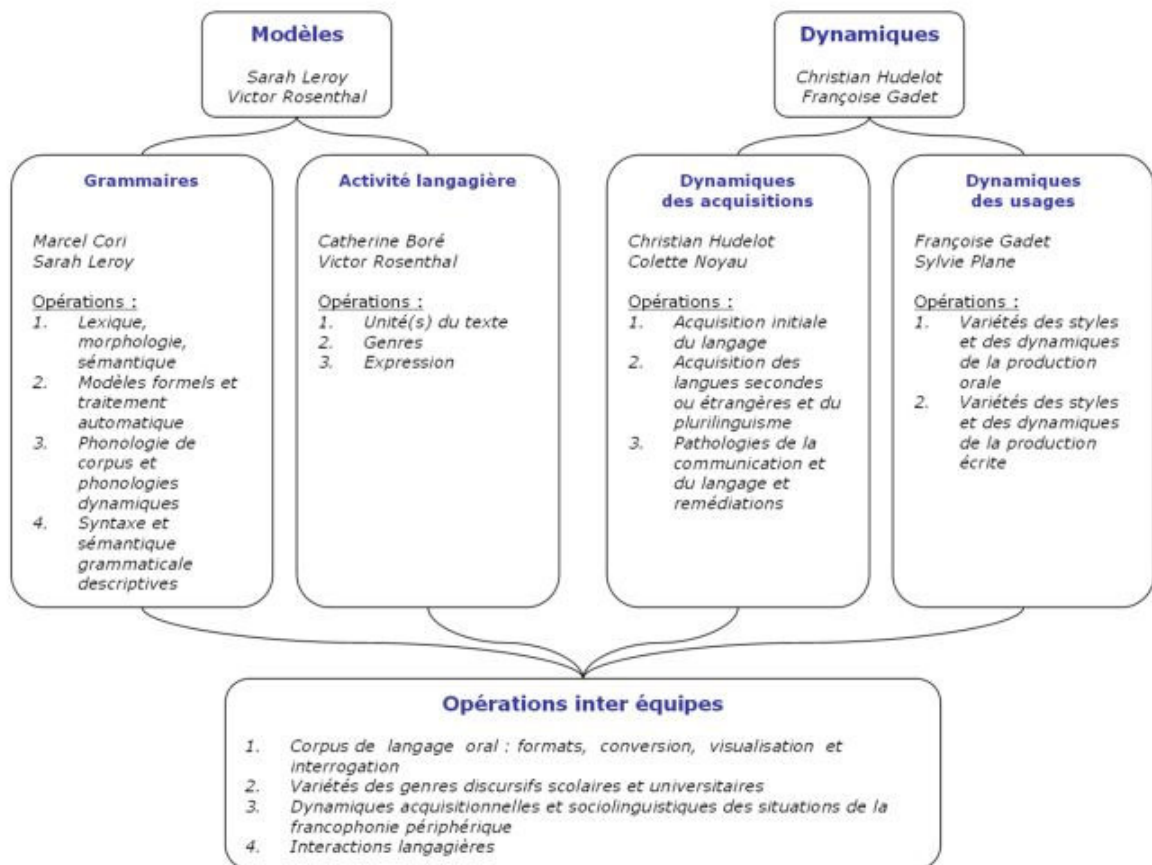
### ***Informatique sur projet (J.-L. Minel, A. Tchobanov, R. Walter)***

\* Etudes, développement et déploiement d'outils informatiques dans le cadre de projets de recherche du laboratoire.

\* Développement et maintenance d'applications informatiques spécifiques en linguistique, dans le respect des normes et des standards en vigueur (gestion de corpus).

\* Support méthodologique dans le domaine de la linguistique textuelle.

### **3.2.2. ORGANIGRAMME SCIENTIFIQUE**



### 3.2.3. RECAPITULATIF QUANTITATIF POUR LE PROJET 2009-2012

<b>Statut (2009-2012)</b>	<b>Nombre</b>
<b>Chercheurs</b>	<b>11</b>
<i>dont CNRS</i>	8
<i>dont INSERM</i>	3
<i>dont émérite</i>	1
<b>ITA (CNRS)</b>	<b>8</b>
<i>ITA permanents</i>	6
<i>ITA CNRS en CDD</i>	2
<i>dont ITA gestion et administration</i>	2
<b>Enseignants-chercheurs</b>	<b>46</b>
<i>dont professeur associé à temps partiel</i>	1
<i>dont émérites ou honoraires</i>	4
<b>Total membres rattachés à titre principal</b>	<b>65</b>

<b>Statut au 1/09/07</b>	<b>Nombre</b>
<b>Doctorants</b>	<b>119</b>
<i>dont ITA en CDD</i>	2
<b>Jeunes chercheurs associés</b>	<b>12</b>
<i>dont ITA en CDD</i>	2

<b>Statut au 1/09/07</b>	<b>Nombre</b>
<b>Membres associés</b>	<b>47</b>

### **3.3. PROJET SCIENTIFIQUE - PERIODE 2009-2012**

#### **Avertissements**

- Le statut des personnes mentionnées est leur statut au 1/09/2007.
- Les chercheurs et les enseignants-chercheurs qui sont mentionnés ont tous rempli une fiche individuelle d'activité et ont ainsi fait part de leur souhait de participer au projet scientifique pour les années 2009-2012. L'organigramme et certains aspects du projet scientifique sont bien évidemment susceptibles d'évoluer en fonction des départs à la retraite effectifs, les demandes qui seront accordées d'éméritat, etc.
- Les références bibliographiques ayant servi à élaborer le projet ne sont pas mentionnées.
- Les doctorants indiqués dans les opérations sont des doctorants inscrits en 2006-2007.
- Nous ne mentionnons pas dans chaque opération le détail de thèses en cours ; ces informations sont disponibles dans la partie « Bilan ».

### 3.3.1. COMPOSANTE *MODELES*

#### **Coordination : Sarah Leroy & Victor Rosenthal**

Naguère réservés à des secteurs directement justiciables d'une formalisation mathématique, les modèles font aujourd'hui partie du paysage habituel des sciences humaines et sociales. Non pas qu'une formalisation à marche forcée ait radicalement transformé la nature de ces sciences, mais parce qu'avec la modélisation se sont ouvertes des perspectives d'intelligibilité, notamment à l'égard des phénomènes complexes, récalcitrants aux méthodes descriptives, qui paraissaient inimaginables il y a peu. Les sciences du langage ont été parmi les premières à anticiper et accompagner ce tournant épistémologique ; elles ont d'ailleurs eu la sagesse de ne pas s'enfermer dans une voie dogmatique, et ont très tôt reconnu l'intérêt de se doter d'une large panoplie de modèles (descriptifs, statistiques, logico-algébriques, neuro-mimétiques, génétiques, morphodynamiques, systèmes complexes).

Avec ces nouveaux outils théoriques et méthodologiques, les recherches linguistiques se sont rapidement ouvertes à des problématiques « connexes » (gestes co-verbaux, formes sémiotiques, normativité, mimesis, origine du langage, langage intérieur, catégorisation, langage et perception, déterminisme culturel, etc.), qui intéressent également d'autres sciences (sciences cognitives, psychologie culturelle, anthropologie sémiotique, sociologie), mais dont aucune ne peut prétendre les affronter seule. Cette interdisciplinarité nouvelle, exigée par l'objet, mobilisant effectivement les acquis de plusieurs disciplines, dans un cadre théorique conçu d'emblée pour constituer et étudier les objets sur ce mode foncièrement interdisciplinaire (pour ses descriptions, méthodologies et modèles) est en train de transformer profondément les façons de faire la linguistique.

Au-delà donc de la modélisation proprement dite, le secteur *Modèles* du laboratoire *MoDyCo* s'inscrit résolument dans cette évolution des recherches linguistiques : diversité des modèles, ouverture thématique et interdisciplinarité.

Partant des aspects internes de la linguistique (lexique, morphologie, phonologie, sémantique), l'équipe *Grammaires* se consacre à l'élaboration de modèles syntaxiques formels dans une perspective de traitement automatique, à l'établissement de modélisations dynamiques à partir de corpus de données phonologiques, à la représentation du sens dans la construction lexicale, et à la recherche des corrélations entre la forme et le sens dans de grands corpus du français.

L'équipe *Activité langagière* étudie les conduites langagières dans le contexte d'interaction sociale, où les faits langagiers sont comme traces de conduites humaines socialement contextualisées. On s'intéresse à la langue comme constituée par l'ensemble des textes qui l'attestent et qui forment conjointement une unité signifiante. On revient sur la disposition expressive de l'expérience humaine qui, pour une part, fournit à l'activité langagière sa motivation et son « modèle dynamique », pour retracer la phylogenèse des langues humaines ou retrouver le modèle « archaïque » de l'expressivité dans l'inscription corporelle du langage (et du geste qui lui est affilié).

### 3.3.1.1. Equipe *Grammaires*

**Coordination : Marcel Cori & Sarah Leroy**

#### Membres de l'équipe *Grammaires*

<b>Membres rattachés à titre principal</b>	Chercheurs enseignants-chercheurs et	K. Baschung, M. Cori, S. David, C. de Loupy, G. Desagulier, M. Desmets, D. Flament, P. Gréa, S. Kahane, M. Klein, A. Lacheret, B. Laks, D. Le Pesant, D. Leeman, S. Leroy, F. Muller, A. Sores	17	22
	ITA	G. Mallet (aussi doctorante), A. Tchobanov	2	
	Membres émérites ou honoraires	A. Balibar-Mrabti, A. Delaveau, F. Kerleroux	3	
<b>Doctorants</b>	A. Azari Nadjaf, E. Bestchastnova, O. Bondeelle, E. Cheung, C. Faivre, T. Geerts, P. Gourdet, C. Graml, P. Harnay, Y. Homma, E. Jouili, A.G. Konte, A. Lablanche, M. Lureau, G.-M. Mallet (aussi ITA CDD), L. N'Kondo, C. O'Brien, K. Ostby, G. P. Renello, C. Roger, E. Saint-Jean, M.-P. Sales, S. Saulnier, C. Salinas-Kahloul, O. Taheri, P. Tomatis		26	
<b>Jeunes chercheurs associés</b>	F. Delumeau, I. Khammari, B. Lavieu, H. Ounis, C. Vaguer.		5	
<b>Membres associés</b>	J.-C. Anscombe (CNRS), J.-P. Bruckert (Paris 10), C. Garitte (Paris 10), P.-P. Haillet (Cergy-Pontoise), J.-P. Jaffré (CNRS), M. Kahloul (Gabès, Tunisie), P. Larrivée (Aston), P. Lauwers (Leuven), R.-J. Lavie, C. Sanguin-Bruckert (Paris 10), F. Villoing (Paris 8).		11	
<b>Collaborations extérieures</b>	M. Amar (BPI), E. d'Andrade (Lisbonne), A. Auchlin (Genève), M.-J. Béguelin (Neuchâtel), G. Bergounioux (Orléans), A. Berrendonner (Fribourg), P.-M. Bertinetto (Pise), P. Blumenthal (Cologne), J. Brandao de Carvalho (Paris 8), B. Calderone (Pise), J.-P. Chevrot (Grenoble), G. Cislaru (Paris 3), J. Coleman (Oxford), L. Degand, (Louvain la Neuve), S. Detey (Rouen), A. Dister (Louvain la Neuve), J. Dubois (Paris 10), F. Dubois-Charlier (Aix-Marseille 1), J. Durand (Toulouse), J. François (Caen), K. Gerdes (Paris 3), G. Girard (Paris 3), J.-P. Goldman (Genève), J. Goldsmith (Chicago), U. Gut (Augsburg), P. Hadermann (Gand), G. Hassler (Potsdam), M. Ihadjadene (Paris 10), O. Inkova (Genève), S. Izral (Tel Aviv), H. Jacobs (Nimègue), C. Kelling (Konstanz), G. Kristofersen (Bergen), F. Lareau (Montréal et Université Paris 7), M. Lecolle (Metz), S.-H. Lee (Séoul), D. Legallois (Caen), J. Léon (CNRS), S. Loiseau (Limsi), C. Lyche (Oslo), J.-M. Marandin (CNRS), J. Meinschäfer (Konstanz), I. Mel'cuk (Montréal), J.-M. Merle (Aix-Marseille I), P. Mertens (Louvain), A. Mettouchi (Nantes), E. Nardi (Rome), N. Nguyen (Aix-Marseille 1), C. Onguene-Essono (Yaoundé), J.-C. Pagès (Baho), C. Pagliano (Nice), M.-G. Pak (Suwon), R. Panckhurst (Montpellier 3), P. Péroz (Metz), M. Pierrard (Bruxelles), A. Polguère (Montréal), N. Quayle (Lille), M. Silberztein (Besançon), A.-C. Simon, (Louvain la Neuve), A. Tifrit (Nantes), D. Tribout (Paris 7), D. van Raemdoncq (Bruxelles), M. von Oostendorp (Amsterdam), S. Wauquier-Gravelines (Paris 8), H.-Y. Yoo (Paris 7).		64	

#### Présentation générale de la problématique scientifique

L'équipe *Grammaires* recouvre un ensemble de travaux en cours et de projets à développer qui ont pour spécificité, au sein du laboratoire *MoDyCo*, de s'intéresser aux aspects internes

de la linguistique dans une visée descriptive et modélisatrice (ce qui la situe dans le secteur *Modèles*) et de s'intéresser à ses objets d'étude en premier lieu en ce qu'ils forment des systèmes linguistiques.

La langue étudiée est le plus fréquemment le français, comme c'est naturel dans un laboratoire de linguistique française, mais on ne s'interdit pas pour autant d'aborder d'autres langues (anglais, allemand, basque, breton, hébreu, hongrois, inuit, kabyle, polonais), ne serait-ce que dans une visée comparative, ou pour étendre ou tester les capacités descriptives et prédictives d'un modèle donné. Les compétences multiples de l'équipe permettent de n'ignorer aucune dimension « interne » de la langue et d'en aborder la phonologie et la prosodie, la syntaxe et la morphologie, la sémantique, ainsi que, naturellement, les rapports entre ces différents niveaux d'analyse complémentaires (prosodie et syntaxe, morphologie et sémantique, etc.).

Du point de vue des méthodes, le travail de l'équipe *Grammaire* se caractérise par un va-et-vient permanent entre données et modèles, ainsi que par une méta-réflexion sur la nature des unes et des autres. Ainsi, les analyses s'appuient toujours sur des descriptions fines et réalistes des phénomènes linguistiques considérés, pour lesquelles on peut utiliser de grands corpus, des données rassemblées *ad hoc* ou des constructions et des tests fondés sur la compétence du locuteur. Ces pratiques ne vont pas sans une réflexion sur les observables (on s'interroge sur le statut à accorder aux investigations sur corpus, selon la nature des corpus) et la mise à disposition d'une partie (qui est appelée à augmenter encore) de ces observables à destination de la communauté scientifique d'une part et, d'autre part, une participation à la réalisation d'outils de traitement des corpus, ainsi qu'à l'évaluation de ces outils. Parallèlement, on développe, tout en les confrontant aux données et à leurs descriptions, des modèles, qui relèvent essentiellement du formalisme à base logico-algébrique et du cognitivisme, deux approches actuelles qui s'inscrivent pleinement dans le contexte international de la recherche dans les domaines concernés. Ces modèles se déclinent en plusieurs options théoriques non exclusives mais plutôt complémentaires, dont on trouvera le détail dans les pages qui suivent.

L'équipe *Grammaires* se structure en quatre opérations majeures :

L'opération *Lexique, morphologie, sémantique* a pour objet d'étude l'ensemble des unités lexicales, envisagées du point de vue de leurs propriétés internes, morphologiques et sémantiques. Il s'agit avant tout de décrire les opérations de construction lexicale, sur les plans formel, catégoriel et sémantique et d'établir des généralisations permettant d'aboutir à une représentation raisonnée du lexique incluant les mots simples comme les mots construits. Morphologique et sémantique allant de pair, l'opération s'enrichit d'une dimension sémantique plus marquée que précédemment, ce qui permettra d'aller plus loin dans la représentation du sens dans la construction lexicale et dans la description de la structure du lexique (avec de nouvelles catégories descriptives, issues de l'approche cognitive, pour le traitement de certains types lexicaux), de confronter des approches sémantiques diverses à des problématiques lexicales complexes (allomorphie, polysémie, lacune lexicale), mais aussi d'augmenter l'empan des réalisations considérées, en y incluant les constructions.

L'opération *Modèles formels et traitement automatique* se consacre d'une part à l'élaboration de modèles syntaxiques formels dans une perspective de traitement automatique, d'autre part à la question de l'utilisation d'instruments (corpus et outils) informatiques en linguistique. Ainsi, un travail de formalisation sera mené sur certains aspects purement syntaxiques



(prédication verbale, ellipse) ou syntactico-sémantiques (scalarité et manière dans le domaine de la comparaison) ; la formalisation permettra également de lier intonation et syntaxe et de comprendre leurs relations. Les travaux de modélisation auront des retombées sur le plan de l'outillage (analyse linguistique multi niveaux, extraction d'informations sur le web, segmentation automatique de corpus) comme sur le plan épistémologique du rôle et de la place des corpus dans le travail du linguiste.

L'opération *Phonologie de corpus et Phonologies dynamiques* vise à développer, à partir de données rassemblées en corpus, les bases descriptives de l'usage phonologique et prosodique contemporain et d'établir, à partir de ces descriptions et pour ces données, les modélisations dynamiques les plus propres à en rendre compte dans toute leur diversité et toute leur variation. Ainsi, la constitution du corpus PFC sera poursuivie et exploitée pour l'analyse phonologique de la liaison et du 'e' muet, de la prosodie dans ses relations avec ces réalisations phonologiques et pour la modélisation de la syllabe. Cette opération est donc active dans la production, l'annotation / étiquetage (problèmes méthodologiques et épistémologiques inclus) et la mise à disposition de données orales, mais exploite aussi ces données orales pour des descriptions réalistes et des modélisations théoriques, sans oublier une réflexion historique et épistémologique faisant retour sur la discipline elle-même.

L'opération *Syntaxe et sémantique grammaticale descriptives* s'organise autour de la recherche des corrélations entre la forme et le sens. Elle s'appuie spécialement sur l'exploitation de grands corpus du français (emplois des verbes ou locutions prépositionnelles du français), mais également sur l'analyse comparative de langues. Ainsi, il s'agit par exemple de systématiser les propriétés observées dans un corpus de verbes français, ou d'avancer dans le recensement des locutions prépositionnelles du français, de décrire comment elles sont constituées et comment elles s'articulent pour former des constructions syntaxiques. L'objectif est notamment la traduction, et aussi la traduction automatique. Il s'agit enfin de valider (ou d'invalidier) des hypothèses relatives à l'énonciation en étudiant la grammaticalisation des marques d'attitude énonciative (comme l'ironie) dans certaines langues.

Cette équipe, articulée autour des disciplines fondamentales de la linguistique synchronique, bénéficie d'un certain nombre d'atouts qui lui permettront de développer une recherche de qualité dans le champ français et international.

Harmonieusement composée de « poids lourds » de la discipline et de plus jeunes (enseignants-) chercheurs, elle échappe ainsi à la menace démographique qui plane sur la plupart des laboratoires français. Cette équipe conjugue diversité des approches et concentration des moyens dans une perspective collaborative. En raison de la complémentarité des compétences et des approches, différents types de théories sont étudiés et confrontés, de même que les différentes dimensions de la langue sont constamment considérées en parallèle. Les étapes, de la description à la modélisation, sont conçues comme complémentaires. L'équipe s'inscrit ainsi dans les tendances actuelles des recherches en linguistique, où les programmes, de longue haleine, sont nécessairement multidimensionnels et collectifs. D'un point de vue purement scientifique, l'équipe participe au renouveau des approches linguistiques, marquées par une récente réorganisation des plans théorique et descriptif, sans pour autant aborder avec naïveté les données et les approches bottom-up : si les recherches sont souvent menées sur des corpus, on n'oublie pas de bien considérer les limites de cette approche et de ne jamais les disjoindre de la modélisation visée. Bien sûr,

l'insertion dans la communauté scientifique passant entre autres par la participation à la constitution de ressources et à la mise à disposition de ces ressources (pour leur utilisation à d'autres recherches ou comme garanties de la reproductibilité des observations), les membres de l'équipe s'y emploient et une augmentation sensible de ces ressources « ouvertes » est programmée.

Par ailleurs, l'équipe *Grammaires* est très liée à la formation. Sur le plan des cursus proposés à Paris 10, elle est impliquée dans l'enseignement et l'encadrement de plusieurs masters (masters recherche « Langue, usages, modèles » et « Traitement automatique des langues », master professionnel « Documents électroniques et flux d'informations »), dont les séminaires (au niveau du master 2) mettent en contact étudiants de master et doctorants. Bon nombre de projets et stages du master « Documents électroniques et flux d'informations » et du master « Traitement automatique des langues » se font au sein du laboratoire *MoDyCo*, ce qui permet d'insérer les étudiants dans un contexte de recherche scientifique et de préparer, en amont, des thèses. L'encadrement doctoral est aussi une des forces de l'équipe : relativement peu de doctorants (moins de 4 par encadrant), pour un encadrement de qualité, souvent financé (une allocation de recherche et un post-doctorat par an pour le laboratoire, actuellement plusieurs thèses financées sur contrats ANR ou européens), et sur des sujets qui s'inscrivent, et s'inscriront de plus en plus clairement, dans le projet global de l'équipe.

Enfin, l'ouverture de l'équipe sur l'extérieur se fait aussi par son insertion dans des réseaux scientifiques, qu'il s'agisse de collaborations ponctuelles ou régulières en France et dans le monde, d'implications dans des actions de recherche coordonnées (PRC, GDR, etc.) ou dans des contrats de recherche financés. La partie « bilan » de ce rapport fait apparaître la capacité de certaines opérations à répondre à des appels d'offre et à mener à bien des contrats de recherche ; cette partie plus prospective indique déjà plusieurs projets de réseaux, de programmes internationaux ou plus modestement de colloques et de publications collectives, qui montrent que cette activité va se poursuivre et se développer, contribuant ainsi à renforcer la visibilité de l'équipe et de ses membres.

### **Liste des opérations**

- 3.3.1.1.1. Lexique, morphologie, sémantique
- 3.3.1.1.2. Modèles formels et traitement automatique
- 3.3.1.1.3. Phonologie de corpus et phonologies dynamiques
- 3.3.1.1.4. Syntaxe et sémantique grammaticale descriptives

#### **3.3.1.1.1. *Lexique, morphologie, sémantique***

##### **1. Membres de l'opération**

*Membres rattachés à titre principal (7 : 2 PR, 2 MC, 3 CR)*

Françoise Kerleroux (responsable)

Sarah Leroy (responsable)

Didier Bottineau

Sophie David

Guillaume Desagulier

Philippe Gréa

Sylvain Kahane

*Doctorants (3)*

Olivier Bondeelle

Coralie Roger

Sophie Saulnier

*Membres associés (3)*

Pierre Larrivée (Aston University)

René Lavie

Florence Villoing (Université Paris 8)

*Collaborations extérieures (12)*

Georgeta Cislaru (Université Paris 3)

Olga Inkova (Université de Genève, Suisse)

Carmen Kelling (Université de Konstanz)

Michelle Lecolle (Université de Metz)

Dominique Legallois (Université de Caen)

Sylvain Loiseau (LIMSI)

Judith Meinschäfer (Université de Konstanz)

Igor Mel'cuk (Université de Montréal)

Claudine Pagliano (Université de Nice)

Alain Polguère (Université de Montréal)

Delphine Tribout (Université Paris 7)

Dan van Raemdonck (Université de Bruxelles, Belgique)

## **2. Problématique**

Le lexique, traditionnellement situé au carrefour de la sémantique, de la syntaxe et de la morphologie, est composé d'unités qui constituent des tous, tout en restant accessibles par strates, par couplage/découplage de ces strates, selon les fonctions qui sont identifiées par les théories de la morphologie, de la sémantique, ou de la syntaxe. Les unités multidimensionnelles qui constituent le lexique exigent un traitement combinant des approches différentes. Mais la perspective lexicaliste commune dans laquelle nous nous situons consiste à envisager que le maximum d'informations linguistiques est inclus dans les propriétés sémantiques des unités lexicales et que ce sont ces propriétés qui sont projetées dans les deux plans des structures syntaxiques (syntagmes et phrases) et des structures morphologiques (lexèmes complexes). L'objectif est donc de découvrir, savoir classer, factoriser et noter les propriétés sémantiques des unités lexicales. Les dimensions identificatrices des unités lexicales, des lexèmes, ne sont pas traitables au nom d'une correspondance Forme/ Sens, directe ou indirecte. D'où des investigations sur les cas d'allomorphie, de polysémie, d'alternances de constructions.

### **2.a. Actions**

#### **1) Morphologie**

Les descriptions et théorisations qui faisaient de la morphologie une application particulière de règles définies ailleurs, en l'occurrence par les règles de réécriture de la syntaxe, ont été

récusées, et les travaux de recherche depuis 30 ans (Aronoff 1976) définissent la discipline comme un ensemble d'opérations formelles et sémantiques, limitées au domaine du mot, dont les langues font deux usages : la flexion, qui est la réalisation morphologique de la syntaxe, et la dérivation, qui est la réalisation morphologique de la formation de lexèmes. L'enjeu théorique et concret de la recherche dans la discipline est alors de découvrir des généralisations morphologiques qui ne se prêtent à aucun parallèle, ou mise en correspondance avec d'autres constructions linguistiques.

- Divers phénomènes conduisent à concevoir la morphologie dans ses propres termes :

(a) Des unités lexicales propres : les bases non autonomes d'origine grecque ou latine. Ainsi la morphologie prend pour bases de ses opérations de construction, par composition (*parricide, aqueduc, anthropophage*) ou par affixation (*galactique, onirique, incinérer*) des lexèmes qui ne se manifestent comme tels dans aucun syntagme ou phrase.

(b) Unicité ou pluralité des formes phonologiques assignées aux lexèmes : c'est un fait purement morphologique que tel lexème soit équipé de plusieurs radicaux, non assignables à des différences sémantiques ou syntaxiques, qui sont identifiés exclusivement par l'usage qu'en fait la morphologie soit flexionnelle (indexation des radicaux d'un verbe dans l'espace thématique que définit le paradigme de la conjugaison), soit dérivationnelle (identification d'un radical caché, servant de base aux affixations en *-ion, -eur, -if* ; *form(er)* vs *format-* dans *formation, formateur*).

(c) Unicité ou pluralité des contenus sémantiques associés à une forme phonologique. L'identification sémantique des lexèmes doit être rapportée aux constructions (Goldberg, 1995) dans lesquelles ils apparaissent, et la morphologie lexicale met en jeu ces différences ou alternances de construction, sous la forme cruciale de lacunes. Ainsi on identifie les contenus sémantiques distincts associés à des formes uniques selon qu'il s'agit d'un verbe de mouvement (à direction inhérente) ou d'un verbe d'activité, identités lexicales distinctes qui se manifestent dans la dérivation impossible ou possible de noms déverbaux agentifs en *-eur* (*monter, descendre*), ou selon qu'il s'agit d'un verbe de localisation ou non, ce que manifeste la dérivation impossible ou possible de noms déverbaux de procès (*comporter, comprendre*). Dans ces champs de recherche (allomorphie, polysémie), la composante de morphologie se présente comme un outil d'exploration, en particulier parce qu'elle permet de discerner des lacunes dans les constructions morphologiques

- Parallèlement, il est nécessaire de réexaminer la notion de morphème. Si des travaux importants de ces vingt dernières années ont montré son inadéquation, d'autres travaux continuent de l'utiliser et de la promouvoir. Il devient alors intéressant de s'interroger sur les raisons de cette situation, en dehors d'hypothèses relevant de la sociologie de la science. Nous partons des travaux de Kerleroux & Marandin (1994), dans lesquels les traitements proposés des mêmes données (en l'occurrence, le GN « sans nom ») se laissent répartir dans les deux modes de connaissance proposés par Pariente (1973) soit, dans ses termes, la « connaissance par modèle » et la « connaissance par système ». Nous reprenons ce type d'approche pour développer l'hypothèse de Kerleroux (2006), selon laquelle l'approche morphématique tomberait sous le mode de la connaissance par système. Dans un deuxième temps, nous nous interrogerons sur les fondements « idéologiques » du mode de connaissance par système, « idéologique » faisant ici référence aux travaux de S.J. Gould (1982), qui ont montré comment des logiques, des valeurs, des cadres de référence non scientifiques traversent la pensée scientifique.

## **2) Sémantique**

L'action Sémantique a pour objectif 1) d'interroger la représentation sémantique des mots construits et en particulier la place du sémantique dans les processus morphologiques, de façon à obtenir une représentation du sens des mots construits en relation avec celui des mots simples, sans oublier le cas du nom propre 2) de développer, vérifier et/ou interroger la validité de théories cognitives fonctionnelles, essentiellement à partir de données du français. Du point de vue méthodologique, cette action donne une importance centrale aux corpus et aux outils de fouille.

### **2a) Types sémantiques**

Il s'agit d'étudier la structure du lexique, du point de vue des liens entre les unités lexicales, liens à la fois paradigmatiques (polysémie, synonymie, antonymie, hyperonymie, etc.) et syntagmatiques (cooccurrence lexicale restreinte) et de celui de l'analyse sémantique des lexies. La perspective choisie est celle du typage des lexèmes, à la fois au niveau de l'élaboration de la représentation formalisée des sens des lexies et de l'étude du rôle des types sémantiques dans la structuration du lexique, mais également à celui de la formalisation des liens lexicaux et de la modélisation dans le cadre de l'élaboration de lexiques formels.

Dans un premier axe on s'intéressera aux types « objet » et « propriété » (Godard & Jayez, 1996), en particulier les noms d'objets dénotant des humains (N-*eur* en français) et à la possibilité d'utiliser l'hypothèse sémantique des types lexicaux à la catégorie du nom propre (phénomènes de changement ou d'empilement de types, comme les toponymes événementiels). Il s'agira enfin d'élaborer des outils et des méthodes pour noter la représentation sémantique des lexèmes, dans un travail de comparaison à partir de Jackendoff (1990), Wierzbicka (1996), Pustejovsky (1995), Beard (1995), Lieber (2004).

Dans un second axe, l'étude de la polysémie, en particulier la classification des différents types de liens de polysémie, sera centrale du fait de son caractère incontournable dans l'identification des lexies. La modélisation des liens lexicaux non standard (ne relevant pas d'un type répertorié comme l'antonymie, l'intensification, etc.) viendra la compléter.

### **2b) Grammaires de constructions et stock lexical**

Il s'agit d'observer le rôle et l'importance des constructions, au sens des grammaires de constructions (Langacker (1987), Lakoff (1987), Fillmore *et al.* (1988), Goldberg (1995), Croft (2001)), qui s'appuient sur le principe selon lequel les séquences linguistiques linéaires qui fondent toute communication verbale finissent par faire émerger des schémas langagiers réguliers plus ou moins figés qui prennent la forme de constructions grammaticales, allant du simple morphème à la proposition la plus générale, qui ont toutes en commun d'être des assemblages symboliques de forme et de sens. Cette approche permet d'éclairer le rôle de la mémoire, de l'apprentissage, et la part de créativité, dans la construction des unités de la grammaire, sans pour autant devoir opérer une distinction, factice au vu du grand nombre de constructions à mi-chemin entre le lexical et le procédural, entre lexique et syntaxe.

À la lumière de cette approche, on privilégiera l'observation des classes ou types lexicaux qui interviennent dans des constructions données (type d'adjectifs dans les comparaisons à parangon Adj + GN) et l'influence de la construction sur les unités grammaticales qui s'y inscrivent. Cette perspective permettra aussi de revisiter la notion de polysémie, ainsi que les phénomènes de changement et de fixation dans la grammaire des locuteurs. Différentes classes de phénomènes seront abordées, dont la notion générale d'expression du plus haut degré, à travers l'étude de comparaisons intensives, et l'analyse de constructions émergentes du français et de l'anglais.

## **2c) Imagerie, profilage et saillance**

Il s'agit de développer la notion d'imagerie issue de la grammaire cognitive, qui permet de distinguer des unités quasi-synonymes (unités grammaticales : déterminants (les indéfinis « vagues » *quelques* et *plusieurs*) et prépositions) par des différences de profilage. On envisage d'étendre la caractérisation en termes de profilage et de saillance à la classe des noms, en particulier les noms collectifs (*forêt, bouquet*) et de collection (*argenterie, vaisselle*) et les massifs pluriels (*représailles, haricots*). Ces différentes classes présentent en effet des similitudes formelles intéressantes que le concept de profilage peut aider à systématiser. Ces questions entretiennent en outre des rapports étroits avec la question du pluriel, spécifiquement traitée dans le cadre cognitif. De plus, une analyse comparative des approches qui ont abordé ces données contribuera à éclairer le champ : ainsi, les notions de discret, dense et compact liées à l'approche énonciative peuvent être comparées à celles d'entité, de massif et de massif répliqué dans la grammaire cognitive ou aux notions de groupe et de somme en sémantique formelle.

## **2.b. Résultats attendus et perspectives**

On souhaite développer, sur la base du séminaire de morphologie, déjà bien installé, et de celui de sémantique, qui a démarré en 2006, un séminaire « Lexique », qui conjoindrait l'un et l'autre, ressemblerait les forces et permettrait des discussions approfondies et des rencontres suivies, de façon à développer les collaborations entre morphologues et sémanticiens.

L'opération verra deux de ses doctorants soutenir leur thèse d'ici le prochain contrat (C. Roger et S. Saulnier). Certains de ses membres participeront à l'organisation du Deuxième colloque international de l'AFLiCo (2009).

Enfin, deux HDR sont en préparation, celle de P. Gréa, sur la place des modèles cognitifs et interprétatifs dans la représentation du sens lexical, et celle de S. Leroy, sur les aspects linguistiques du processus cognitif général de l'évaluation orientée vers le plus haut degré.

### **3.3.1.1.2. Modèles formels et traitement automatique**

#### **1. Membres de l'opération**

*Membres rattachés à titre principal (9 : 4 PR, 2 MC, 1 PAST, 1 CR, 1 IR)*

Marcel Cori (responsable)

Sylvain Kahane (responsable)

Karine Baschung

Sophie David

Annie Delaveau

Marianne Desmets

Anne Lacheret

Claude de Loupy

Jean-Luc Minel

*Doctorants (3)*

Eliane Cheung

Anne Lablanche

Marie-Pierre Sales

*Jeune chercheur associé (1)*

Fabrice Delumeau

*Collaborations extérieures (17)*

Muriel Amar (BPI)

Antoine Auchlin (Université de Genève)

Marie-José Béguelin (Université de Neuchâtel)

Alain Berrendonner (Université de Fribourg)

Liesbeth Degand (Université de Louvain la Neuve)

Anne Dister (Université de Louvain la Neuve)

Kim Gerdes (Université Paris 3)

Majid Ihadjadene (Université Paris 10)

Shlomo Izral (Université de Tel Aviv)

François Lareau (Université de Montréal et Université Paris 7)

Jacqueline Léon (CNRS)

Jean-Marie Marandin (Université Paris 7)

Piet Mertens (Université de Louvain)

Amina Mettouchi (Université de Nantes)

Rachel Panckhurst (Université Montpellier 3)

Alain Polguère (Université de Montréal)

Hi-Yon Yoo (Université Paris 7)

## **2. Problématique**

Ces dernières années, dans le domaine du Traitement automatique des langues (TAL), les méthodes dites de « TAL robuste », fondées principalement sur des calculs statistiques effectués à partir de grands corpus, ont pris une importance grandissante. Ceci n'invalide pas une recherche en TAL qui s'appuie sur une approche linguistique et une modélisation formelle. En effet on s'aperçoit que les méthodes de TAL robuste produisent des réalisations imparfaites, qu'on essaie de corriger par la prise en compte de connaissances linguistiques. Et la modélisation formelle est indispensable si l'on souhaite une linguistique falsifiable et la réalisation d'outils informatiques raisonnés et révisables.

Par ailleurs, les recherches en linguistique se nourrissent de plus en plus d'investigations effectuées sur de grands corpus à l'aide d'outils informatiques. Il est important d'étudier les conséquences de cette utilisation des corpus sur les résultats de la recherche et, en particulier, d'observer si le choix des corpus et des outils de recherche a une influence sur les descriptions obtenues. Inversement, on s'intéressera aux outils effectivement utilisés en TAL, afin notamment de contribuer à la réalisation d'outils efficaces, d'avoir un jugement de linguistes sur ces outils, de participer à une évaluation des outils.

### **2.a. Actions**

#### **1) Description linguistique dans le cadre de modèles formels, comparaison de ces modèles**

Nous travaillons au développement de différents modèles formels, ce qui permet, à travers une confrontation permanente, une mise en question et un enrichissement de chacun des modèles. La description des faits linguistiques se fait, dans la mesure du possible, indépendamment du choix du modèle que l'on a retenu, ainsi peut-on distinguer ce qui dépend effectivement des modèles. Dans la période à venir divers axes sont envisagés :

(1) La description sera centrée sur le lexique des classes fermées (articles, connecteurs, pronoms et tous les adverbes au comportement atypique). Il existe aujourd'hui des lexiques des classes ouvertes et semi-ouvertes (verbes, noms, adjectifs, adverbes, prépositions, déterminants) de qualité acceptable. Le développement d'une grammaire formelle, une fois résolues des questions comme l'ordre des mots ou l'accord, consiste essentiellement à modéliser le comportement des mots outils de la langue. Par exemple, l'intégration d'un mot comme *fois* qui possède simultanément un comportement de nom (il se combine avec un déterminant) et une distribution d'adverbe (*il est venu plusieurs fois*) peut remettre en cause l'architecture complète d'un modèle. Un fragment de grammaire de français en Grammaire d'Unification Sens-Texte est en cours de développement.

(2) Dans le cadre des grammaires d'arbres polychromes, outre les descriptions de l'infinitif et de l'adjectif en français, il est prévu d'avancer le travail qui a été effectué sur le créole guadeloupéen. En particulier, la combinaison des auxiliaires de prédication sera représentée sous la forme de relations qui s'intègrent dans une base de données et on pense en déduire une procédure originale d'analyse morphosyntaxique.

(3) Dans le cadre des recherches en HPSG, on développera un travail de formalisation des descriptions en cours dans le domaine des constructions comparatives (voir collaborations Desmets & Moline à ce sujet). On proposera une analyse syntaxique formelle des comparatives de degrés et des comparatives dites de manière (en *comme*), on se penchera sur les problèmes d'interface syntaxe-sémantique, en particulier sur les problèmes de congruences entre analyse syntaxique et analyse sémantique posés par les formes à ellipses et l'on proposera une théorie de l'ellipse inspirée de Ginzburg & Sag (2000) pour le traitement des fragments phrastiques.

## **2) Intonosyntaxe**

La problématique explorée est la suivante : la modélisation automatique et la compréhension du rôle des marqueurs prosodiques et syntaxiques dans l'organisation du discours (ANR corpus en cours d'évaluation). En pratique, il s'agit de développer des outils de segmentation automatique du discours (unités minimales du discours) basés sur trois types d'indices : exclusivement prosodiques, exclusivement syntaxiques, intonosyntaxiques pour d'une part proposer une évaluation quantifiée de chaque type d'indices dans la segmentation de la chaîne parlée, d'autre part améliorer les performances des analyseurs syntaxiques des langues en se fondant sur la prosodie.

## **3) Observation des corpus, outillage informatique et évaluation**

A la suite du travail de Cori et David (à paraître), il s'agit tout d'abord de s'interroger sur les rapports entre l'attesté (ce qu'on trouve dans les corpus) et le « possible de langue » (Milner), et en déduire ce qu'il en ressort pour des théories linguistiques fondées sur les corpus. Ce qui dépend évidemment des supports sur lesquels on va chercher ses données. Ainsi, si on se réfère au Web, on trouvera abondamment des productions « hors norme », dont il faudra savoir quel statut leur donner ; mais de telles productions se trouvent aussi dans des corpus plus orthodoxes. Dès lors, le problème se pose d'un réexamen de ce que l'on peut considérer comme étant la norme linguistique. Inversement, on s'interrogera sur des productions qui paraissent à l'évidence appartenir à la langue mais qui sont absentes de certains corpus, si grands soient-ils.



On participera également à la réalisation d'outils informatiques, parmi lesquels des outils d'exploration des corpus, des outils d'extraction d'informations, notamment d'extraction d'informations lexicales à partir de corpus ou du Web, des outils de transformation des informations à travers des grammaires de texte, ainsi qu'à la réalisation d'un analyseur linguistique multi niveaux. Par ailleurs, le travail déjà engagé sur l'évaluation sera poursuivi.

## **2.b. Résultats attendus et perspectives**

### *1) Description linguistique dans le cadre de modèles formels, comparaison de ces modèles*

Les résultats attendus consistent essentiellement en l'écriture de fragments de grammaires dans les différents modèles considérés. En GUST seront intégrées un certain nombre de connaissances lexicales, ainsi des connaissances sur les adverbes et les locutions adverbiales. En GAP, on attend une grammaire de l'adjectif ainsi que de l'infinitif en français. En HPSG la grammaire devra intégrer une interface syntaxe-sémantique portant sur les comparatives. Par ailleurs la perspective est de faire entrer ces grammaires ou fragments de grammaires dans des analyseurs effectifs, qui seront construits ou adaptés d'analyseurs existants.

### *2) Intonosyntaxe*

Les résultats attendus sont les suivants :

A) Pour le français : a) outils pour l'élaboration et l'analyse discursive d'un corpus de référence en français parlé basé sur l'annotation intonosyntaxique semi-automatique, b) analyseur syntaxique automatique avec balises discursives intégré au système de synthèse de la parole développé à l'IRCAM.

B) Perspective inter langues : utiliser tout ou partie des outils développés sur des langues typologiquement non apparentées (pour l'heure : polonais, berbère, hébreux) et y voir plus clair concernant les points suivants : (1) quels sont les traits intonosyntaxiques invariants dans la segmentation de la parole humaine, (2) dans quelle mesure les outils de traitement automatique développés pour le français peuvent-ils être également appliqués à des langues à tradition orale (berbère), (3) quelle est la limite inter langues des ressources informatiques proposées (pour l'heure : travail sur le polonais et l'hébreux).

### *3) Observation des corpus, outillage informatique et évaluation*

Il s'agit :

- de produire une analyse effective et une classification de ce que l'on peut qualifier comme phénomènes hors normes dans les corpus. Ceci pourra donner lieu à une thèse.
- de construire effectivement et d'évaluer les outils informatiques nécessaires.

## **3.3.1.1.3. Phonologie de corpus et phonologies dynamiques**

### **1. Membres de l'opération**

*Membres rattachés à titre principal (5 : 2 PR, 1 MC, 1 IR, 1 IE)*

Anne Lacheret (responsable)

Bernard Laks (responsable)

Marc Klein

Atanas Tchobanov

Richard Walter

*Doctorants (15 dont 5 en cotutelle)*

Allahverdi Azari Nadjaf

Evguenia Bestchastnova

Twan Geerts  
Carolin Graml  
Maxence Lureau  
Géraldine Mallet  
Laurent Nkondo  
Katherine Ostby  
Gian Paolo Renello  
Omid Reza Taheri  
Patrick Tomatis

*Chercheur associé (1)*

René Lavie

*Collaborations extérieures (20)*

Ernesto d'Andrade (Université de Lisbonne)  
Gabriel Bergounioux (Université d'Orléans)  
Pier-Marco Bertinetto (SNS Pise)  
Joaquim Brandao de Carvalho (Université Paris 8)  
Basilio Calderone (SNS Pise)  
Jean-Pierre Chevrot (Université de Grenoble)  
John Coleman (Université d'Oxford)  
Sylvain Detey (Université de Rouen)  
Jacques Durand (Université de Toulouse)  
Jean-Philippe Goldman (Université de Genève)  
John Goldsmith (Université de Chicago)  
Ulrike Gut (Université d'Augsburg)  
Haïke Jacobs (Université de Nimègue)  
Geert Kristofersen (Université de Bergen)  
Chantal Lyche (Université d'Oslo)  
Noël Nguyen (Université de Provence)  
Marc von Oostendorp (Meertens Instituut)  
Anne-Catherine Simon (Université de Louvain la Neuve)  
Ali Tifrit (Université de Nantes)  
Sophie Wauquier-Gravelines (Université Paris 8)

## **2. Problématique**

Bien que pratiquement inaperçue en France où les phonologies post-génératives classiques sont restées dominantes, et où même les phonologies optimalistes ne se sont que très peu installées, le champ conceptuel de la phonologie contemporaine internationale a connu des changements récents très substantiels. La relation entre théorie et données, modèles abstraits et observables s'est totalement inversée. De modèles sous-déterminés par les données, on est passé à des approches qui placent au premier plan les observables et les phénoménologies construites explicitement. Les approches quantitatives, statistiques et probabilistes se sont développées. Les grammaires de construction qui modifient profondément la conception que l'on peut avoir du stockage lexical, les modèles occurrencealistes qui accordent une importance centrale à la structure statistique des données, les modèles exemplaristes et leurs effet sur la conception de l'apprentissage, ont tous conjointement contribué à redonner aux

données d'observation la place première qui étaient la leur dans les approches classiques. Dans ce contexte théorique international renouvelé, les opérations de phonologie du laboratoire *MoDyCo* visent à définir, modéliser et promouvoir, *une phonologie de corpus*. Chacune des actions de l'opération vise, dans le cadre d'une phonologie de corpus, à développer les bases descriptives de l'usage phonologique contemporain. Il s'agit également d'inscrire ce travail sur les observables dans le débat théorique actuel concernant les modélisations dynamiques les plus propres à rendre compte des usages et des variations qui les marquent. La modélisation et l'analyse critique des théories en concurrence sont donc toujours menées de pair avec l'analyse et la description des phénoménologies phonologiques du français contemporain.

## **2.a. Actions**

### **1) Participation au programme coopératif *Phonologie du français contemporain***

Avec ce programme, dirigé par Jacques Durand, Chantal Lyche et Bernard Laks, financé par l'ANR (2005-2007) et la DGLFLF (2002-2007), il s'agit de poursuivre et d'achever le travail d'enquête de terrain, dans l'ensemble du monde francophone afin d'atteindre l'objectif de 500 locuteurs et de 800 heures d'enregistrement. Le travail de numérisation, d'analyse et de codage contextuel permettant d'intégrer ces données à la base actuelle sera poursuivi. Un effort particulier sera fait pour développer l'accessibilité à ces données via la plateforme internet, les outils de fouille et d'interrogation en ligne. Du point de vue de l'équipe *MoDyCo*, ses acquis en la matière lui confèrent une qualité de leader et d'organisateur au plan européen (cf. *infra* le réseau Corpho).

### **2) Participation au programme coopératif *Prosodie du Français contemporain***

Ce programme est dirigé par Chantal Lyche et Anne Lacheret. L'action 2 s'inscrit à l'interface de l'action 3 ci-dessous et de l'opération Modèles formels et traitement automatique (action « intonosyntaxe »). La recherche conduite a pour objectif d'utiliser les indices suprasegmentaux dans la modélisation des phénomènes de liaison et de 'e' muet. Elle doit permettre de répondre aux deux questions suivantes : (1) dans quelle mesure l'élision d'un 'e' laisse des traces prosodiques ? (2) de quelle façon la structure prosodique conditionne-t-elle en partie la prononciation ou l'élision du 'e' en français ? Plus concrètement, il s'agit de développer une grille de codage des données prosodiques en interaction avec le niveau segmental, d'automatiser autant que possible la tâche de codage, et enfin de développer des outils informatiques de requêtes, de visualisation et d'analyse des résultats pour le test et la falsification des hypothèses que nous avons pu formuler, et ce à deux niveaux d'analyse : 1° relation entre phonologie de surface et phonologie profonde, 2° contraintes informationnelles sur les constructions segmentales et suprasegmentales. Cette action est fortement reliée à d'autres travaux de l'équipe, dans la mesure où d'une part une partie des développements informatiques ici conçus sont communs à l'action « intonosyntaxe », d'autre part les données PFC prosodiquement codées viennent nourrir les analyses à l'interface de l'intonation, de la syntaxe et de la pragmatique, ainsi que la modélisation automatique qui en résulte. À l'inverse, les traitements syntaxiques automatiques envisagés dans l'opération Modèles formels et traitement automatique auront des retombées immédiates ici.

### **3) Modélisations phonologiques de la liaison, du ‘e’ muet et des phénomènes connexes en phonologie du français contemporain**

Cette action s’inscrit dans le projet « De la phonologie aux formes lexicales : liaison et cognition en français », dirigé par Jacques Durand et financé par l’ANR (2008-2010). Il s’agit, s’appuyant sur les acquis de PFC comme sur ceux de l’analyse critique des modèles, de faire retour sur la liaison, le ‘e’ muet et les phénomènes connexes en phonologie du français. Les modèles exemplaristes et les grammaires de construction sont plus particulièrement mis à contribution, dans des travaux d’analyse phonologique qui explorent l’interface phonie/graphie, les modes de stockage des constructions et l’acquisition de ces phénomènes phonologiques variables.

### **4) Analyse dynamique de la syllabe à l’aide de modélisations connexionnistes**

Dans la ligne des phonologies dynamiques et des analyse de la syllabation proposées par Goldsmith puis Laks, il s’agit de construire des modèles neuromimétiques de la syllabation. L’utilisation des oscillateurs temporels et de la détection des synchronisations de décharge sera poursuivie. Cette modélisation dynamique vise à expliquer la syllabe et la syllabation comme étant le produit de l’intégration linéaire des segments. L’intégration des éléments sonores dans la chaîne résulte en une auto organisation dont la syllabe constitue la forme résultante.

### **5) Histoire et épistémologie des sciences cognitives et de la phonologie au 20<sup>e</sup> siècle**

Cette action a pour objectif de revenir sur l’histoire et l’épistémologie de la phonologie et de principaux concepts de la linguistique du 20<sup>e</sup> siècle (langue parole, compétence performance, structure et système, usages et corpus, grammaire universelle et machine dérivationnelle etc.) en les éclairant par une histoire parallèle de la logique mathématique, de l’intelligence artificielle et de la théorie de l’information.

### **6) Réseau européen Corpho (Corpora in Phonological Research)**

Cette action participe au réseau Corpho dirigé par Geert Kristofersen (Université de Bergen), Marc von Oostendorp (Meertens Instituut) et Ullrike Gut (Université d’Augsburg), et financé par l’ESF (2006-2007)<sup>1</sup>. Il s’agit d’animer un groupe européen pour favoriser les convergences dans la technologie des corpus oraux, les codages et étiquetages (en relation avec la TEI).

## **2.b. Résultats attendus et perspectives**

D’un point de vue général, cette opération permettra l’enrichissement des ressources mises à disposition sur le site, en particulier par la distribution en ligne d’outils pour traiter et analyser les ressources PFC. On prévoit aussi la mise au point de formats pour l’annotation et la lecture de données phonologiques dans une perspective d’interopérabilité et d’échange.

1) Action 1 : La synthèse de l’opération PFC sera marquée par la publication de 3 volumes chez Hermès en 2008-2012. Comme au cours des deux années précédentes, un colloque international et une réunion nationale sont organisés chaque année. Deux thèses seront soutenues en 2007 (Twan Geerts et Géraldine Mallet) et trois thèses au cours du contrat (Evguenia Bestchastnova, Carolin Graml et Katherine Ostby).

---

<sup>1</sup> <http://www.corpho.eu/>

Un important projet d'utilisation systématique de PFC pour l'acquisition de la phonologie du Français langue seconde est actuellement mis en place par Sylvain Detey (Université de Rouen) Jacques Durand, Chantal Lyche et Bernard Laks. Il coordonne des enseignants-chercheurs de FLE/FLS dans une quinzaine de pays (voir les pages dédiées de <http://www.projet-pfc.net>). Ce dernier projet fait l'objet de plusieurs demandes de financement en cours d'évaluation au près de la DGLFLF, du Ministère de la recherche et des universités, de l'Agence Mondiale de la Francophonie et l'Association des Universités Francophones. Un volume est accepté pour publication chez Ophrys en 2009.

2) Action 3 : Le projet ANR sur lequel s'appuie l'action 3 démarre en 2008. Sont prévus une rencontre nationale et un colloque international chaque année. Les résultats paraîtront dans deux volumes collectifs projetés, le premier consacré à l'analyse phonologique de la liaison et du 'e' muet, le second consacré à la divergence phonie/graphie et au question d'apprentissage normal ou pathologique qui en découlent.

3) Action 4 : Cette action est menée en collaboration avec l'École Normale Supérieure de Pise dans le cadre d'un projet franco-italien. Deux colloques, l'un à Paris et l'autre à Pise, sont organisés chaque année. Deux thèses seront soutenues au cours du contrat (Patrick Tomatis et Omid Reza Taheri)

4) Action 5 : Il s'agit de terminer le livre *Battles in the Mind Fields* dont la première moitié à été écrite lors du séjour de Goldsmith au laboratoire.

5) Action 6 : Sont prévues la publication en 2009 d'un *Handbook of Speech Corpora and Corpus Phonology* et l'organisation d'une université européenne d'été consacrée aux corpus oraux et à la phonologie de corpus à l'Université d'Augsburg en 2008.

#### **3.3.1.1.4. Syntaxe et sémantique grammaticale descriptives**

##### **1. Membres de l'opération**

*Membres rattachés à titre principal (10 : 5 PR, 4 MC, 1 CR)*

Danielle Leeman (responsable)

Antoinette Balibar-Mrabti

Annie Bertin

Didier Bottineau

Danièle Cogis

Danièle Flament

Denis Le Pesant

Claire Martinot

François Muller

Anna Sörös

*Doctorants (10 dont 6 en cotutelle)*

Carole Faivre

Julie Glikman

Patrice Gourdet

Patrick Harnay

Yukiyo Homma

Estelle Jouili

Amadou Konte

Jean-Christophe O'Brien

Erika Saint-Jean

Claudine Salinas-Kahloul

*Jeunes chercheurs associés (4)*

Ichraf Khammari

Belinda Lavieu

Houda Ounis

Céline Vaguer

*Membres associés (9)*

Jean-Claude Anscombe (CNRS)

Jean-Pierre Bruckert (Université Paris 10)

Catherine Garitte (Université Paris 10)

Pierre Patrick Haillet (Université de Cergy-Pontoise)

Jean-Pierre Jaffré (CNRS)

Mongi Kahloul (Université de Gabès, Tunisie)

Peter Lauwers (Université de Leuven, Belgique)

René-Joseph Lavie

Christine Sanguin-Bruckert (Université Paris 10)

*Collaborations extérieures (15)*

Peter Blumenthal (Université de Cologne)

Jean Dubois (Université Paris 10)

Françoise Dubois-Charlier (Université de Provence)

Jacques François (Université de Caen)

Geneviève Girard (Université Paris 3)

Gerda Hassler (Université de Potsdam)

Seong-Heon Lee (Université de Séoul)

Jean-Marie Merle (Université de Provence)

Emma Nardi (Université de Rome 3)

Christine Onguene-Essono (Université de Yaoundé)

Jean-Christophe Pagès (Baho)

Man-Ghyu Pak (Université de Suwon, Corée du sud)

Pierre Péroz (IUFM de Metz)

Nigel Quayle (Université de Lille)

Max Silberztein (Université de Besançon)

## **2. Problématique**

Les points communs aux chercheurs ici rassemblés, par-delà d'éventuelles différences théoriques, sont méthodologiques : (a) une procédure abductive concluant à la plausibilité d'une hypothèse à partir d'un ensemble d'observations interprété en fonction d'un savoir commun, (b) par conséquent une méthode reposant sur l'observation des formes et de leurs propriétés, dont le résultat descriptif est interprété (c) selon quelques grands présupposés comme le principe de naturalité (corrélation forme/sens) ou le principe de pertinence du sémiologique, ou (c') selon une théorie de référence qui s'en trouve validée ou falsifiée. Le travail porte sur de grands corpus du français (inventaire et classement de tous les emplois des verbes ou locutions prépositionnelles du français) et/ou l'analyse comparative de langues diverses, soit dans la perspective traductologique (locutions françaises et allemandes,

anglaises, espagnoles, néerlandaises), soit dans la perspective plus théorique de validation ou invalidation d'hypothèses de linguistique générale (français, basque, breton, espagnol, inuit...).

L'objectif est d'organiser les propriétés observables afin de déterminer (par hypothèse) la structuration sémantique d'un système et d'établir une corrélation entre la forme et le sens. Les thématiques retenues sont le verbe (Action 1), les locutions (Action 2), les marques d'attitude énonciative (Action 3). Dans les trois cas sont prises en compte différentes langues, que l'intérêt se porte plus spécifiquement sur la traduction ou la traduction automatique, la comparaison ou la typologie, la pertinence translinguistique des concepts grammaticaux (comme « le sujet »), la vérification d'une théorie linguistique générale, la didactique des langues.

## **2.a. Actions**

### **1) Le classement syntactico-sémantique des emplois des verbes français**

A partir du postulat que la langue est un système autonome dont la structuration sémantique ne peut par conséquent être assimilée à aucun ordre autre que le sien, l'objectif est d'explicitier, compléter et systématiser les emplois et propriétés inventoriés dans le corpus Les Verbes français (mis à disposition de tous via le site *MoDyCo*) afin de construire la structuration sémantique propre à cet ensemble lexical. L'exploitation de ce corpus a donné lieu à une première publication dans *Langue française* (n°153, mars 2007) et à une demande de financement ANR dont nous attendons l'expertise finale.

Une sous-action *Didactique de la syntaxe et du lexique* se fixe l'objectif de constituer à partir de cette base un outil à la portée des enseignants de français langue maternelle et de français langue étrangère : les grandes lignes en sont présentées dans *Orthographe : innovation et pratiques de classe*, à paraître en décembre 2007 aux Presses Universitaires de Namur (Belgique). Dans le domaine de la didactique du français comme langue étrangère, le programme Hubert Curien franco-coréen STAR (financé par Égide) implique un partenariat avec la Corée.

Une sous-action *Dictionnaires électroniques* se fixe d'autre part l'objectif d'optimiser, à partir du corpus des verbes, les performances d'un analyseur syntaxique (définition du projet ANR ci-dessus mentionné).

### **2) Les locutions prépositionnelles et prépositives du français, traduction et traduction automatique**

L'objectif est d'inventorier les locutions du français et leurs emplois (premières données mises à la disposition de tous via le site *MoDyCo*), d'en décrire les propriétés internes (constitution) et externes (positions syntaxiques) à partir de corpus attestés et de procéder à leur(s) traduction(s). Cette mise au jour permet d'améliorer les résultats actuels en matière de traduction automatique aussi bien que dans l'analyse automatique des langues. Ce programme, défini en janvier 2007 (cf. les actes dans *Modèles linguistiques* 55 : 2007 : 1), a débuté par la liste des locutions en sous, décrites et analysées sur les plans syntaxique, morphologique, sémantique, en ce qui concerne le français (journée d'étude à Paris 10 le 6 juillet 2007) et traduites en allemand, anglais, néerlandais (journée d'étude à Aix-en-Provence le 27 octobre 2007).

Ce projet, sous l'intitulé « Locutions, langues, cultures, cognition », constitue le programme d'un Collège doctoral franco-allemand Paris 10-Potsdam 2007-2009, accepté pour financement par l'UFA (Université franco-allemande) en avril 2007. Dirigé, pour la partie française, par Danielle Leeman, il comporte en outre, au titre de *MoDyCo*, une équipe de base

formée de : Annie Bertin, Danièle Flament-Boistrancourt, Denis Le Pesant, François Muller, et, pour ce qui est des chercheurs associés, Jean-Claude Anscombre (CNRS), Peter Lauwers (Leuven, Belgique). Il implique, toujours du côté français, quatre doctorants en co-tutelle avec Potsdam : Estelle Jouili, Julie Glikman, Christophe O'Brien, Erika Saint-Jean – liste susceptible de s'accroître selon les inscriptions en thèse à venir.

### **3) La diversité des langues dans la représentation de l'universel, le cas des marques d'attitude énonciative**

Étant admis le postulat que l'universel s'instancie dans les langues, de manière grammaticalisée (syntaxe, morphologie) ou non (il s'agit alors de déterminer comment), l'hypothèse que l'on se donne à démontrer est que la grammaticalisation des marques d'attitude énonciative (comme l'ironie) dans certaines langues valide la théorie polyphonique de l'énonciation.

#### **2.b. Résultats attendus et perspectives**

*Soutenances de thèse et d'HDR* : les 7 thèses inscrites en 2005 sont suffisamment avancées pour venir à soutenance en 2008 (E. Jouili, Y. Homma, P. Gourdet, A. Conte) et 2009 (J.-C. O'Brien, E. Saint-Jean, P. Harnay). D. Bottineau et D. Cogis ont une HDR en préparation qui viendra à soutenance avant le fin du prochain contrat quadriennal.

1) Action 1 : *Le classement syntactico-sémantique des emplois des verbes français* fait l'objet d'un projet ADONIS (candidature soumise fin 2007) – l'objectif subsistera même si la candidature n'est pas sélectionnée par le CNRS (mais sa réalisation sera évidemment plus longue). La sous-action *Didactique de la syntaxe et du lexique* donnera lieu à la publication d'un ouvrage diffusant les résultats du programme de l'Action précédente en les adaptant au public des enseignants. Le projet franco-coréen Hubert Curien financé par Egide implique l'organisation de deux colloques dans les deux années à venir, l'un à Séoul, l'autre à Paris 10 (dont les actes seront publiés respectivement en 2009 et 2010) et la publication des résultats pour le public des enseignants des deux langues comme langues étrangères. La sous-action *Dictionnaires électroniques* a fait l'objet d'un projet ANR (soumis en mars 2007) - l'objectif subsistera même si le dossier n'est pas retenu (auquel cas la réalisation du programme sera évidemment plus longue).

2) Action 2 : Le travail entrepris en 2007 avec les locutions en *sous* se poursuit (le numéro de *l'Information grammaticale* de mars 2008 ne pourra témoigner de l'ensemble des résultats, qui donneront lieu à publication sur un autre support) et s'étend aux autres locutions selon la même démarche : journées trisannuelles alternativement à Aix-en-Provence et Paris 10, permettant de comparer les résultats individuels et de préparer la publication de leurs actes. Le collègue doctoral bi-national franco-allemand articulé à ce projet suppose lui-même des séminaires, tantôt à Potsdam, tantôt à Paris 10. Le travail sur le français à partir de Dubois et Dubois fait l'objet d'un projet ADONIS (candidature soumise fin septembre 2007).

3) Action 3 : le mémoire d'HDR d'Anna Sores (soutenance le 8 novembre 2007) programme un ouvrage à paraître (2009) sur la comparaison des langues et les concepts interlangues. Les premiers résultats de D. Bottineau et R.-J. Lavie donnent lieu à un article à paraître dans *Langue française* en 2009.



### 3.3.1.2. Equipe *Activité langagière*

**Coordination : Catherine Boré & Victor Rosenthal**

#### **Membres de l'équipe *Activité langagière***

<b>Membres rattachés à titre principal</b>	Chercheurs enseignants-chercheurs et	H. Bays, A. Bertin, C. Boré, D. Bottineau, S. Bouquet, J.-J. Briu, M.-L. Elalouf, C. Facq-Mellet, G. Gargiulo, J.-F. Jeandillou, F. Rastier, V. Rosenthal, A. Trévisé	13	18
	ITA	J. Couto, J.-L. Minel, R. Walter, S. Weiser (aussi doctorante)	4	
	Membres émérites ou honoraires	M. Arrivé	1	
<b>Doctorants</b>	M. Ait Saada, L. Constant, A. Dhoukar, O. Djedid, K. Elkandoussi, K. H. Kang, B. Karamifar, W. Kurawan, L. Mansour, E. Mogollon Ballestas, A. Muller, F. Nore, A. Rosa, Y. Schochenmeier, T. Thumrongsanta, S. Tyaglova, A. Vraciu, S. Weiser (aussi ITA CDD), S.-J. Woo		19	
<b>Jeunes chercheurs associés</b>	F. Amardeilh, V. Muni Toke		2	
<b>Membres associés</b>	D. Ablali, C. Camugli Gallardo, H.-T. Kim, R.-J. Lavie, D. Malrieu, F. Sitri, Y.-M. Visetti, I. Weill		8	
<b>Collaborations extérieures</b>	C. Cornejo Alarcon, P. Bisiacchi, P. Cadiot, M. de Fornel, J. Lassègue, J.-L. Petit, D. Roulland, L. Lundquist, D. Wonsever		9	

#### **Présentation générale de la problématique scientifique**

L'équipe « activité langagière » trouve son identité thématique et théorique dans une conception du langage comme une totalité langue-parole qui, en tant qu'unité de langue-parole, est exercice de cette faculté. C'est en ce sens que nous pouvons entendre d'abord « l'activité » : à la fois comme potentialité dynamique de la langue encore non réalisée, et comme réalisation dans la parole.

Les références théoriques sont alors, de façon classique, celles de Saussure – en particulier celui des *Eléments de linguistique Générale* présenté par Bouquet, 2002- et Guillaume, ainsi que le courant néo-saussurien de Hjelmslev.

Cette conception justifie l'existence de différents projets se situant dans une linguistique de la parole, partageant une notion de la langue non pas « une et unique », mais constituée par l'ensemble des textes qui l'attestent, textes oraux ou écrits, qui constituent l'unité d'analyse, en tant qu'unité signifiante.

La question de l'*unité* ou des *unités* du « *texte* » (au sens pluriel que l'on vient d'évoquer), est alors posée : elle interroge le principe de compositionnalité, habituel en linguistique, et la discrétisation des objets linguistiques ; ceci revient à poser la question du continu et du discontinu dans la langue, dans une confrontation entre différents modèles théoriques (théories « cognitives », traitement automatique, approches grammaticales, théories interprétatives du sens, etc.).

Le langage est une institution. En tant que pratique instituée, il peut être saisi au travers de conduites humaines sémiotisées dont les textes constituent l'une des manifestations. C'est dans ce sens spécifique qu'on pourrait encore entendre l'« activité langagière » (Bronckart) : des conduites langagières élaborées dans l'interaction sociale, en interdépendance avec leur contexte d'élaboration. Les faits langagiers sont alors saisis comme traces de conduites humaines socialement contextualisées. Cette position est par ailleurs cohérente avec une approche non essentialiste, historique et moniste (relation physique-psychique) du langage. La langue s'élaborant et évoluant, comme tout système sémiotique, au sein de pratiques sociales différenciées à l'intérieur de domaines et de champs de pratiques, le *genre* apparaît comme le lieu où s'articulent les contextes d'énonciation et les normes langagières, le lieu où s'explicitent les jeux de différentes sémiotiques qui informent le texte et contraignent les dimensions cognitives de l'interprétation.

Dans la mesure où l'activité langagière vise la parole et les physionomies affiliées (gestes, postures, regard, mimiques faciales), les recherches s'ouvrent aux linguistiques cognitives et aux courants dynamiques et 'culturalisants' des sciences cognitives.

L'activité langagière tire sa motivation et son « modèle dynamique » d'une disposition expressive (participative) de l'expérience humaine, qui est le mode premier de gestion de l'intersubjectivité et le lieu d'union originnaire du perceptif, de l'émotionnel, du motivationnel, de l'axiologique, du cognitif et du culturel. Le concept d'expression est donc essentiel pour retracer la phylogenèse des langues humaines en même temps que l'émergence et l'évolution de diverses formes et activités symboliques, pour saisir le statut du geste co-verbal, ou celui du langage (même intérieur) comme le lieu où la pensée prend forme et se révèle à la conscience. On retrouve d'ailleurs le modèle « archaïque » de l'expressivité dans l'inscription corporelle du langage, que ce soit sur le plan des propriétés sensorimotrices ou des phases et dynamiques de l'activité de la production de la parole, ou sur le plan sémantique (ex. caractère synesthésique des principales formes adjectivales, transposition sémiotique du corps à l'ensemble du monde sensible).

Les projets de l'équipe, reprenant les termes de cette problématique, s'articulent autour de trois opérations : *Unité(s) du texte*, *Genres*, *Expression*, chaque axe déclinant ses propres théories et méthodes.

### **Liste des opérations**

3.3.1.2.1. *Unité(s) du texte*

3.3.1.2.2. *Genres*

3.3.1.2.3. *Expression*

#### **3.3.1.2.1. *Unité(s) du texte***

##### **1. Membres de l'opération**

*Membres rattachés à titre principal (7 : 4 PR, 1 IR, 2 IE)*

Marie-Laure Elalouf (responsable)

Annie Bertin

Jean-Jacques Briu

Jean-Luc Minel

Richard Walter

Anne Trévisé  
Stéphanie Weiser (aussi doctorante)

*Membre associé (2)*  
Catherine Camugli Gallardo  
Isabelle Weill

*Doctorants (11 dont 1 en cotutelle)*  
Mekioussa Ait Saada  
Laetitia Constant  
Khadija Elkandoussi  
Banafsheh Karamifar  
Eugenia Mogollon Ballestas  
Agnès Muller  
Aiala Rosa  
Svetlana Tyaglova  
Alexandra Vraciu  
Stéphanie Weiser (aussi ITA CDD sur un contrat ANR)

## **2. Problématique**

La conception du langage dans laquelle s'inscrit l'équipe « activité langagière » pose la double question épistémologique des unités et de l'unité du texte, c'est-à-dire des relations entre le local et le global, le statique et le dynamique dans la construction de l'interprétation. Cette question traverse aujourd'hui de nombreuses théories du texte (cf. le numéro 163 de la revue *Langages* (sept. 2006), dirigé par D. Legallois). Dans le cadre de la présente opération, la solidarité du continu et du discontinu sera traitée sous trois angles : la représentation des unités du texte dans le cadre de son traitement informatique, leur transformation en diachronie et leur dénomination en didactique des langues.

L'adoption par la communauté scientifique d'une représentation structurée des textes, conçue sous forme d'éléments imbriqués (TEI) soulève des problèmes théoriques et méthodologiques qui sont travaillés dans l'action « Représentation et traitement informatique des unités du texte ». Une attention particulière est portée aux différents niveaux de relations entre les unités textuelles. Les différents projets qui relèvent de cette action visent une modélisation du discours, avec des applications notamment dans les domaines de la compréhension et de l'extraction d'informations.

En diachronie, il s'agira de s'interroger sur la disparition de certaines unités dans des séries grammaticales et sur l'apparition d'unités résultant d'un processus de grammaticalisation ou de création. La catégorie de la conjonction constitue un poste d'observation privilégié à plusieurs titres : catégorie relationnelle par excellence, elle a subi un renouvellement important entre l'ancien et le moyen français en même temps qu'évoluaient les frontières et les propriétés syntaxiques des unités mises en relation. Il en résulte un double problème terminologique qui rejoint les questions posées en didactique : qu'entend-on par conjonction selon que l'on décrit tel ou tel état de la langue ? et comment dénommer les unités textuelles reliées par les conjonctions ?

En didactique des langues, le poids de la tradition logico-grammaticale a longtemps occulté la question théorique des « unités du texte », et celle, indissociable, de leur dénomination. Aujourd'hui, le foisonnement terminologique impose un examen qui ne dissocie pas les perspectives linguistiques et didactiques : en fonction de quels critères faut-il recourir à des termes spécifiques ou peut-on emprunter à une terminologie en vigueur pour la description de la phrase ? Comment articuler les différents niveaux d'analyse dans une perspective didactique ?

Ces questions renvoient au choix des cadres théoriques permettant de mettre en relation des expressions linguistiques et des activités langagières dans un contexte de communication. Posant conjointement la question des « paliers de traitement » de l'énoncé et des unités correspondant à telle ou telle opération, les grammaires fonctionnelles (Van Vallin & La Polla, 1997) justifient la hiérarchisation des niveaux d'analyse non de façon uniquement interne mais par la relation qu'ils entretiennent avec le discours et la cognition dans une perspective communicative. Il y a là, selon B. Combettes (2005), une perspective pertinente pour une grammaire de phrase renouvelée, compatible avec l'étude du texte et du discours. Dans le cadre de l'action, *Terminologie et dénomination des « unités du texte »*, cette hypothèse sera soumise à la confrontation de linguistes et didacticiens de différentes langues, travaillant dans différents cadres théoriques.

## **2.a. Actions**

### **1) Terminologie et dénomination des « unités du texte »**

Un rapide tour d'horizon fait apparaître les limites et les contradictions des solutions retenues aujourd'hui en didactique des langues. Dans l'enseignement du français dit langue maternelle, la grammaire de phrase relève toujours des concepts et des méthodes du distributionnalisme, tandis que des notions empruntées à la linguistique du texte et du discours ont été ajoutées sans réorganisation conceptuelle d'ensemble. En didactique des langues en revanche, l'approche communicative dans sa version la plus radicale relègue au second plan l'étude du système de la langue. Dans l'enseignement de l'anglais, l'introduction de la théorie des opérations énonciatives est restée par bien des aspects superficielle. Quant au secteur des langues romanes, si une ébauche d'articulation entre niveaux d'analyse émerge depuis peu pour l'espagnol et le portugais, en italien les choix demeurent flous et peu explicites.

On s'appuiera sur la réflexion déjà engagée pendant le précédent contrat, dans le cadre de l'élaboration d'un glossaire de termes linguistiques pour la didactique des langues. L'objectif sera de poursuivre la rédaction d'entrées du glossaire autour de la problématique des unités du texte. Il s'agira de fournir des analyses permettant d'éclairer les choix des rédacteurs de programmes, de manuels et des enseignants en didactique des langues. Le choix des entrées privilégiera la recherche d'articulations entre niveaux d'analyse, par exemple en prolongeant le travail déjà engagé sur le groupe nominal par des entrées sur la construction référentielle, l'identification et l'activation d'un référent dans le discours. Le travail pourra ensuite être élargi aux notions de prédication, de classe de prédicats, etc.

### **2) La subordination dans l'histoire du français »**

Le travail articulera deux axes :

L'étude du « **renouvellement des conjonctions** » pour reprendre la formule de Meillet-disparition, création, selon des schémas cognitifs mis en lumière dans le cadre de la grammaticalisation, mais aussi sous l'action de contacts de langues (en particulier le latin

scolastique), en fonction des genres textuels (importance des textes juridiques) la créativité n'étant pas l'apanage de la langue orale spontanée. Le sujet donnera lieu à un groupe de travail en 2007-2008. On partira d'une présentation de B. Kortman, *Adverbial subordinators in the languages of Europe*, qui fait suite aux travaux du groupe « Adverbial Relations, Operators and Connectives » (resp. van der Auwera) du programme de recherche international EUROTYP.

L'étude de la « **conjonction** », notion aussi floue que tenace dans la grammaire occidentale (pour Aristote c'est aussi bien tel mot-outil que l'agencement qui donne l'*Odyssee*), amène à s'interroger sur le statut de la notion même de subordination voire de phrase, particulièrement dans l'optique des textes médiévaux. Les propositions, la phrase sont-elles des unités pertinentes pour étudier la phrase médiévale ? Cette question rejoint certainement les interrogations d'autres linguistes de l'équipe, en particulier en ce qui concerne les productions scolaires.

A ces deux axes, on associera les travaux suivants :

- Etude d'un préfixe *-re* (non répertorié dans les dictionnaires latins) au sens de « blocage dans la case temporelle considérée » (ex. : *remanere, remanoir* en ancien français, *to remain* en anglais) ;
- *tant* marque d'une opération qui n'a rien à voir avec l'intensité - mais qui est la marque d'entrée dans un monde réduit logique quantifié et connu.

### **3) Représentation et traitement informatique des unités du texte**

L'adoption, par la communauté des chercheurs qui travaillent sur corpus, de standards de balisage (XML) et de ses différences instances dédiées (TEI) soulève, sous couvert de problèmes techniques, de nombreux problèmes théoriques et méthodologiques. Parmi ceux-ci, nous nous intéressons plus spécifiquement au problème de la représentation informatique d'un texte, c'est-à-dire aux unités et aux relations qui le composent. Nous disposons actuellement d'une représentation élaborée dans le cadre du projet Navitexte (Couto 2006, Couto & Minel 2007a, 2007b) et mise à l'épreuve dans différentes applications (Battistelli & al. 2007, Mathieu 2006, Lundquist & al. 2006). Les résultats obtenus nous amènent à enrichir ce modèle en nous confrontant à d'autres types de textes et notamment ceux étudiés par les membres de l'équipe, notamment les corpus scolaires, le discours rapporté, et les pages Web dédiées à un domaine.

D'autre part, nous élargissons notre réflexion aux problèmes posés par l'extraction d'informations dans le cadre du Web sémantique. La plupart des approches actuelles ne prennent pas en compte la dimension textuelle et se limite au contexte local dans leur processus de repérage ou d'extraction d'informations. Les travaux menés par F. Amardeilh (2007) et l'élaboration du langage OPALE a démontré toute l'importance de disposer d'une représentation, et d'un langage formel associé, qui soit apte à accueillir les relations syntaxiques, sémantiques et rhétoriques qu'entretiennent les unités d'un même texte.

Cette opération est adossée à trois projets ANR en cours et à trois collaborations internationales :

Par ailleurs, en collaboration avec l'équipe dirigée par Mme Lita Lundquist (*Département de français Institut F.I.R.S.T, Copenhagen Business School, Danemark*), nous développons le logiciel NaviLire, outil d'aide à la compréhension de textes complexes pour des étudiants FLE qui s'appuie sur la plate-forme NaviTexte. L'objectif est de concevoir un outil qui permette à un apprenant de voir et de naviguer dans un texte entre des unités textuelles assurant la cohérence. Cet outil constitue un instrument didactique très performant et

motivant, tant pour l'apprentissage de la lecture que pour l'apprentissage de la production écrite de textes. Une première version de NaviLire est diffusée au Danemark avec l'ouvrage « Tekstkompetence på fremmedsprog » publié par, Forlaget samfundslitteratur. Ce projet est soutenu par l'ambassade de France au Danemark.

## **2.b. Résultats attendus**

1) Action 1 : on s'appuiera sur la réflexion déjà engagée pendant le précédent contrat, dans le cadre de l'élaboration d'un glossaire de termes linguistiques pour la didactique des langues. L'objectif sera de poursuivre la rédaction d'entrées du glossaire autour de la problématique des unités du texte. Il s'agira de fournir des analyses permettant d'éclairer les choix des rédacteurs de programmes, de manuels et des enseignants en didactique des langues. Le choix des entrées privilégiera la recherche d'articulations entre niveaux d'analyse, par exemple en prolongeant le travail déjà engagé sur le groupe nominal par des entrées sur la construction référentielle, l'identification et l'activation d'un référent dans le discours. Le travail pourra ensuite être élargi aux notions de prédication, de classe de prédicats, etc.

En préparation : Elalouf M.-L. (dir.), 2009, « Groupe nominal et expressions référentielles : approches linguistiques et didactiques », Presses universitaires de Namur, collection Dyptique.

2) Action 2 : organisation à l'horizon 2009 d'un colloque « Fondements et méthodes d'une linguistique diachronique ».

3) Action 3

- Un ouvrage de synthèse intitulé « Filtrage sémantique, approches symboliques » sous la direction de Jean-Luc Minel, aux éditions Hermès dans la collection IC2.

- Une extension de la plate-forme de navigation textuelle, NaviTexte, enrichie par un modèle de la représentation des relations entre des unités textuelles et du logiciel NaviLire.

- L'ensemble des livrables du projet Eiffel ANR-RNTL (fin 2009) avec notamment des ressources linguistiques dédiées au repérage des informations spatio-temporelles dans des pages Web (thèse de Stéphanie Weiser).

- Organisation d'un colloque international (en 2009) avec les équipes *Culture et Communication* de l'Université Avignon (EA 5131), le *GRIPIC* de l'Université Paris-Sorbonne (EA 1498) sur la problématique des « écrits d'écran » dans le Web sémantique.

### **3.3.1.2.2. Genres**

#### **1. Membres de l'opération**

*Membres rattachés à titre principal (10 : 3 PR, 6 MC, 1 DR)*

Catherine Boré (responsable)

Caroline Facq-Mellet (responsable)

Michel Arrivé

Hillary Bays

Annie Bertin

Simon Bouquet

Marie-Laure Elalouf

Giuseppe Gargiulo

Jean-François Jeandillou

François Rastier

*Jeune chercheur associé (1)*

Valelia Muni Toke

*Membres associés (4)*

Driss Ablali (Université de Franche-Comté)

Hui-Teak Kim (Université nationale de Séoul)

Denise Malrieu (CNRS)

Frédérique Sitri (Université Paris 10 et équipe Syled)

*Doctorants (6 dont 1 en cotutelle)*

Ourdia Djedid

Wanrudee Kurawan

Léda Mansour

Tissawas Thumrongsanta

Yevgeny Schochenmeier

Sun-Joo Woo

*Collaborations internes avec l'équipe « Dynamiques des usages »*

## **2. Problématique**

Le genre est un concept déterminant pour les sciences du langage. Il est au cœur d'une réflexion épistémologique sur la complémentarité et l'articulation entre linguistique de la langue et linguistique de la parole. Caractérisé de manière minimale et provisoire par un faisceau de régularités engendrées par des pratiques diversifiées qui le constituent, le genre, tout en répondant à une description linguistique, intègre une dimension traditionnellement reléguée à l'extra-linguistique, identifié généralement par le terme de « contexte ». Cette dimension ne s'entend pas cependant exclusivement dans le sens prioritairement idéologique de l'analyse du discours. Le genre est en effet, par la lignée des textes qui l'attestent, ce qui, d'abord, tient lieu de contexte. Il constitue de cette façon une détermination essentielle du sens, si l'on entend par là, avec Rastier, que le global détermine le local, c'est-à-dire les composants locaux des textes.

L'acception de « textes » se réfère au principe hjelmslevien énoncé dans les *Prolégomènes à une théorie du langage* : « La théorie du langage s'intéresse à des textes(...) » et prend acte du fait que la linguistique selon Hjelmslev part du texte, donnée de base de l'analyse linguistique, chaîne syntagmatique comprenant tous les textes présents passés à venir et appartenant à toutes les langues, à partir de laquelle le linguiste déduit les composantes d'une théorie du langage. Le texte ne s'oppose pas aux textes, ce sont deux aspects du même objet. C'est donc l'objet « textes » qui est le domaine empirique d'analyse et d'observation des genres.

Les différentes actions menées au sein de l'opération développent et approfondissent de manière complémentaire cette conception générale du genre :

1) L'action « Recherches épistémologiques pour une linguistique des genres » élabore le cadre théorique de l'opération.

2) L'action « Caractérisation des genres » permet d'étayer la proposition bakhtinienne en faveur d'une linguistique des genres diversifiée. Dans ce sens, elle est le prolongement empirique de l'action 1).

Le détail des projets souligne la part égale prise par l'approche morpho-syntaxique de l'analyse automatisée des genres et celle des contextes dans la modélisation.

Il s'agit tout à la fois d'explorer des genres liés à des pratiques variées -genres scolaires, genres parlementaires, genres poétiques, genres électroniques, genres métalinguistiques (dictionnaires) etc., et de participer à une réflexion sur la définition et le mode d'appréhension des genres.

3) L'action « Discours rapportés » adopte cette terminologie usuelle, sans préjuger de cette étiquette, pour aborder les phénomènes complexes de la « représentation de discours autres » (Jacqueline Authier-Revuz). Sa particularité consiste à mettre en évidence de manière prioritaire la manière dont le genre détermine l'élaboration de formes langagières et contraint l'interprétation des énoncés. Elle aborde ce problème dans une perspective pragmatique qui prend en compte de façon explicite et raisonnée le dispositif énonciatif lié au genre textuel, puisque celui-ci définit les contraintes pragmatiques de l'interprétation.

L'une des questions abordées sera d'explorer en quoi une approche qui tient compte du genre textuel peut aider à définir les critères linguistiques des frontières entre discours rapportés. Elle s'appuiera également sur les descriptions diachroniques disponibles (Marnette) pour s'interroger sur l'existence de marques identifiables de ces discours. Les corpus étudiés sont contrastés, en synchronie (corpus de signalements d'enfants en danger, corpus de contes utilisés en milieu scolaire, des romans, un corpus de questions de parlementaires au gouvernement etc.) comme en diachronie (littérature d'Ancien Français).

4) Enfin l'action « réécritures de la langue médiévale », à partir de l'alignement des corpus anciens, se pose la question de la normativité des corpus, qui est une préoccupation générique centrale.

### **1) Recherches épistémologiques pour une linguistique des genres**

Cette action est la poursuite d'une recherche en cours. Elle postule qu'une philosophie du langage est directement productive d'une épistémologie linguistique.

A partir de *l'Herméneutique* de Schleiermacher, des derniers travaux de Bakhtine et de l'œuvre du « second » Wittgenstein, on compare les concepts de « genre » chez Schleiermacher et Bakhtine et de « jeu de langage » chez Wittgenstein dans la perspective du programme saussurien. Les travaux empiriques concernent, entre autres, une grammaire de la parole du français oral et des textes littéraires, notamment poétiques.

### **2) Caractérisation des genres**

Les analyses morpho-syntaxiques contrastives sur corpus étendus de différents genres textuels, à visée soit classificatoire soit attributive, se sont multipliées ces dernières années.

Comme dans toute description, le jeu des variables choisies est lié à des variables intermédiaires hypothétiques. Les travaux de F. Rastier et D. Malrieu (2002) ont fait l'hypothèse qu'aux différentes postures qui caractérisent les domaines de pratiques, correspondent des modalités de construction du sens hétérogènes qui renvoient à des textualités différentes et auxquelles pourraient correspondre des configurations préférentielles de modalités énonciatives, de combinaisons modaux / personnes, de constructions de subordination vs juxtapositions, de modes de prédication de détermination (définie / indéfinie), de valeurs sémantiques des temps, d'usage des déictiques, etc. Cette approche



statistique de grandes masses textuelles, permet un repérage de traits stylistiques contrastés selon les genres, à travers la mise en évidence de faisceaux de variables corrélées

Les nouvelles actions envisagent, pour partie, un traitement automatisé de discours scolaires oraux (voir Représentation et traitement informatique des unités du texte » dans l'opération « Unité(s) du texte) et écrits (voir « Genres scolaires » « Echanges électroniques », « genres métalinguistiques), et, pour partie, un traitement manuel (« genres parlementaires »).

## **2a) Genres scolaires**

Ces travaux ont d'abord été empiriques et descriptifs, ils ont visé à approcher la langue dans les manifestations orales et écrites émises par un locuteur spécifique, l'écolier, à travers des observables : les textes (oraux ou écrits) qu'il produit.

Le travail de recueil des données reste fondamental, d'abord parce qu'il n'a jamais été réellement fait au point de constituer une archive en sciences du langage; les textes produits dans l'institution scolaire sont utilisés comme documents sociologiques, historiques ou didactiques, et très peu comme des objets langagiers manifestant l'appropriation de la langue par de jeunes scripteurs, qui renseignent sur la dialectique langue/parole, idiolecte/stéréotype, style individuel d'« erreurs »/style collectif d'« erreurs » ; ensuite parce que les données recueillies sont aussi – nécessairement – l'objet d'une méthodologie et d'une épistémologie.

De ce point de vue, le contexte est primordial : on entend par là sans doute le contexte institutionnel constitué d'éléments extra-linguistiques intériorisés dont les marques se retrouvent dans les textes concrets (Bronckart) ; mais le contexte est aussi composé de textes ; l'intertextualité est le milieu « naturel » des textes, et conditionne l'interprétation sémantique : les textes de référence prescrits et lus à l'école en font partie, de même que la masse des écrits produits dans l'institution. Quelle que soit la manière dont on considère « ces produits » de l'institution, qu'on les appelle « textes » ou « discours », ils relèvent de pratiques qui les inscrivent dans un genre : nous les nommons les *genres scolaires*.

Si la préoccupation (et les références) cognitives ne sont pas premières dans l'énonciation des textes, nous nous référons cependant aux analyses vygotkiennes issues de « Pensée et Langage » : les hypothèses de Vygotski sur la naissance du langage intérieur de l'enfant, prélude de son passage à l'écriture, constituent un élément interprétatif de l'évolution des textes produits.

La prise en compte du « genre » comme catégorie descriptive a déplacé l'analyse vers une réflexion sur la notion de corpus, qui suppose alors de se mobiliser sur le commun autant que sur la variété. Nous avons entrepris un travail de comparaison de textes narratifs appartenant à la vaste famille des « contes » telle qu'elle se manifeste dans les programmes et références institutionnelles de l'école

Il s'agit de numériser, outre des séries homogènes de productions écrites d'élèves rangées sous cette étiquette à l'école, les listes de contes que l'institution prescrit dans les programmes, et sur lesquelles un travail de sériation est tout à fait nécessaire : les contes didactiques, contes littéraires traditionnels, contes traduits, adaptés, réécritures modernes, pastiches etc., sont génériquement apparentés mais le système dans lequel ils apparaissent (isolément ou groupés dans les manuels ou dans les listes) modifie leur sens.

## **2b) Les genres parlementaires**

L'objet de l'étude est triple. Il s'agit d'abord de poursuivre la description des genres de discours liés à l'institution parlementaire, en diachronie et en synchronie. Nous nous attacherons plus particulièrement à l'interpellation. Nous voulons ensuite contribuer à une réflexion méthodologique sur la définition et sur le mode d'appréhension des genres. Enfin,

en cherchant à mettre en évidence la pertinence du concept de genre pour la détermination du sens, nous souhaitons participer à une réflexion plus générale sur une (re)définition d'une linguistique du sens.

### **2c) Echanges électroniques par MSN ou par Chat**

Le travail porte sur deux caractéristiques du genre en émergence susceptibles de le caractériser :

- l'identification d'une « phonologie » iconique des parties visuelles d'une conversation écrite ;
- l'utilisation de la webcam et de l'image en parallèle avec le texte écrit.

### **2d) Genres métadiscursifs**

La recherche a pour cadre le programme d'édition scientifique des écrits linguistiques de Charles Nodier : discours théoriques concernant la motivation iconique, à partir de la *Bibliothèque sacrée grecque-latine* (1826) et de l'*Examen critique des dictionnaires* (1828).

### **3) Discours rapportés**

L'action s'est engagée en 2007 dans une perspective descriptive. Elle tient compte des apports de Marnette, Rosier, qui ont tenté d'ébranler le cadre formel élaboré dans un champ dominé par les analyses de J. Authier-Revuz.

Un traitement automatique par balisage de séquences hétérogènes, commencé sur les romans de M. Duras, sera étendu à une partie du corpus.

Cette méthodologie devrait permettre une analyse fine des dynamiques discursives selon le genre textuel, en particulier les modes d'enchaînements des différents types de discours, leur répartition topologique dans le texte en fonction des propriétés pragmatiques du genre ou du dispositif narratologique.

Une typologie des DR a pu être esquissée, elle part des distinctions suivantes :

- parole au discours direct adressée à autrui (DD) ou à soi-même (monologue intérieur MI), le DD rapporté (DDR) marqué ou non marqué typographiquement, le discours indirect (DI) et monologue intérieur au discours indirect (MII),
- les DR narrativisés : le DRN : parole d'autrui rapportée à la 3S et au temps du récit, avec ou sans incise de dire, et sans introducteur et jonctif ; le DIN : discours indirect narrativisé : on a un *verbum dicendi* à l'imparfait ou PQP; le jonctif peut introduire une suite de phrases faussement autonomes et gardant les marques de temps du DI; le MIN : parole intérieure ou MI narrativisé.

Ces descriptions tiendront compte, dans l'hypothèse d'une modélisation ultérieure, des aspects diachroniques de la représentation de la parole.

Le DR présente en effet dans la littérature en Ancien Français de nombreuses formes non conventionnelles ; il a suscité de nombreux travaux (A. Meiller et S. Marnette, dans la lignée des travaux de B. Cerquiglini).

Plusieurs Bases de données de textes en AF, balisent d'ailleurs le texte selon ce critère (*Le Chevalier à la charrette*, projet de K. Uiti-Sarah Kay à Princeton continué à Poitiers ; l'ensemble des romans de Chrétien de Troyes, projet en cours en collaboration Ottawa-ATILF-Amsterdam, sous la direction de P. Kuntsman).

### **4) Réécritures de la langue médiévale**

L'écriture médiévale, plus encore que d'autres, est très souvent réécriture : par la force des choses, parce que la transmission manuscrite interdit le monolithisme du texte imprimé,

multipliant les variantes qui peuvent faire apparaître la variation diatopique et diachronique ; parce qu'elle se donne dans sa très grande majorité comme la « translation » d'un texte-source, le « conte » ; parce qu'elle se plaît au gré des interpolations, des continuations, des suites, des mises en prose à varier la matière.

On privilégiera deux axes :

- constitution et exploitation de traductions « intralinguales » (ancien français- moyen français) ; le traitement XML allié à un logiciel d'alignement doit permettre des études quantitatives fiables du changement morpho-syntaxique et lexical.
- étude linguistique des mises en prose, qui se développent à la fin du Moyen Age, à partir des romans et des chansons de geste, à partir de corpus informatisés.

## **2.b. Résultats attendus**

1) Action 1 : un colloque : « De la grammaire à l'inconscient : dans les traces de Damourette et Pichon », à Cerisy en 2009 ou 2010, avec Michel Arrivé, Claudine Normand et Valelia Muni Toke.

2) Action 2 : une collaboration (C. Boré, M.-L. Elalouf) s'est constituée avec l'équipe « Dynamiques des usages » sur le projet européen **COST** (genres) « Variétés des genres discursifs scolaire et universitaires » (resp. S. Plane). On se reportera au descriptif de cette opération. Par ailleurs l'HDR de C. Boré, en préparation, est prévue pour 2008-2009.

- Publication de la thèse de Caroline Facq-Mellet sur les questions au Gouvernement. Analyse d'un genre discursif.

3) Actions 3 et 4 :

- A l'initiative de S. Bouquet, le groupe « discours rapportés » fera une synthèse de ses travaux dans un numéro de « Langue Française » en 2009 à l'issue d'un colloque organisé à Paris 10.

- Préparation d'un ouvrage de C. Boré sur le dialogue dans les genres scolaires de fiction  
Diverses journées d'études sont prévues en 2010 sur la linguistique de corpus.

### ***3.3.1.2.3. Expression***

#### **Membres de l'opération**

*Membres rattachés à titre principal (3 : 1 PR ; 2 CR)*

Didier Bottineau (responsable)

Annie Bertin

Victor Rosenthal

*Doctorants (3)*

Kyung-Hwa Kang

Asma Dhoukar

Françoise Nore

*Membres associés (2)*

René-Joseph Lavie

Yves-Marie Visetti (CREA-CNRS, UMR 7656)

*Collaborations extérieures (7)*

Carlos Cornejo Alarcon (Université de Santiago, Chili)

Patricia Bisiacchi (Université de Padoue)  
Pierre Cadiot (Université d'Orléans)  
Michel de Fornel (EHESS)  
Jean Lassègue (CREA, CNRS UMR 7656)  
Jean-Luc Petit (Collège de France - CNRS UMR 7152)  
Daniel Roulland (Université de Haute Bretagne)

## **2. Problématique**

Le terme *expression* dénote ici le régime sémiotique élémentaire de l'expérience qui est en même temps le mode premier de gestion de l'intersubjectivité et le lieu d'union originaire du perceptif, de l'émotionnel, du motivationnel, de l'axiologique, du cognitif et du culturel. Se situant en deçà de l'activité langagière dont elle fournit (en partie) la motivation<sup>2</sup> et à laquelle elle prête sa constitution dynamique, l'expressivité originaire de l'expérience s'incarne tout d'abord dans le caractère physionomique de la perception, où toute forme, toute configuration s'offre comme une tonalité animée, une manifestation spontanée de la vie. La qualité expressive du monde perçu se trouve donc explicitement lisible dans le mode même de présentation des figures – jusqu'à les doter éventuellement d'animation et d'intériorité. Toute expérience subjective, toute perception, est d'emblée qualitative (e.g. chargée d'une valence affective) et sémiotique, son caractère physionomique procédant des dynamiques de constitution des configurations perceptives qui font de l'expérience expression de son propre processus de constitution. Cette constitution dynamique instancie *in fine* la structure fondamentalement participative (et énaïve) de l'expérience humaine.

Une telle couche expressive originaire ne doit pas être prise pour une véritable saisie sémiotique de l'action dans un réseau symbolique, laquelle implique conventionnalisation d'un sens générique et ritualisation des gestes. En effet, l'accès immédiat à la médiation sémiotique ne signifie pas que les formes et les valeurs qui trament l'expérience puissent se concevoir hors du jeu sans fin renouvelé des transactions sociales qui les portent à l'existence. Dans ces jeux socio-sémiotiques, même la part cognitive individuelle se comprend comme perception sémiotique, attention conjointe et participation à une intersubjectivité comportant un vaste répertoire d'interactions ritualisées. Le symbolique émerge alors des pratiques sémiotiques publiques qui structurent l'expérience en domaines d'activités, traçant entre eux des connexions et des homologies, et définissant par là des normes de perception et d'action. L'articulation de l'expressivité « primitive » avec des pratiques sémiotiques publiques permet de concevoir l'émergence de formes et activités symboliques, et à travers elle, celle des langues humaines.

Au palier de l'activité langagière elle-même, la prise en compte de ses dynamiques expressives valorise une perception diachronique du langage comme activité (*enérgeia*) et source de changements, à diverses échelles temporelles, qui vont de la microgenèse de l'acte de parole à l'évolution de la langue au cours de son histoire.

La disposition expressive, à la fois « primitive » et encadré par la socialité du sens, manifeste tant en deçà de l'activité langagière qu'au sein de ses dynamiques de constitution, est récemment devenue l'objet d'intenses recherches et de tentatives de modélisations réunissant sciences cognitives, sciences du langage et sciences sociales. Il s'agit en l'occurrence d'un

---

<sup>2</sup> Difficile en effet de concevoir l'activité langagière sans une disposition expressive en amont de toute velléité de vociférer, serait-ce intérieurement.

véritable tournant théorique et épistémologique par rapport à la tendance très marquée au cours de la deuxième moitié du vingtième siècle dans les sciences du langage comme dans la fédération naissante des sciences cognitives, par ce qui apparaît rétrospectivement constituer un déni des problématiques de l'expression et de la socialité du sens. Soit que les démarches en cause aient prétendu pouvoir ignorer ou différer toute question de sens ou tout motif qualitatif (d'action ou d'expression) ; soit qu'elles les aient réduites à la seule perspective d'un accès à une objectivité indépendante et d'une représentation communicable ; soit enfin que, concédant une certaine socialité du sens (sans pour autant admettre la problématique de l'expressivité de l'expérience), elles aient pensé pouvoir *in fine* la réduire à la recherche d'utilités élémentaires (définies au préalable dans un registre éco-éthologique) et communicationnelles.

Les membres de l'équipe participant à cette opération ont tous contribué ces dernières années aux recherches et à la modélisation des différentes problématiques expressives et diachroniques, que ce soit au niveau de l'expressivité « primitive » de l'expérience et de la sémiotique (la théorie de la microgenèse, Rosenthal, 2004a, 2004b, 2005), de la modélisation des faits d'évolution (Bertin), de l'émergence des langues et du langage (Lassègue *et al.*, 2006, 2007), de l'expression comme lieu où la pensée prend corps et accède à la conscience, des propriétés sensori-motrices des systèmes dans les langues et de la modélisation des articulations sémantiques dérivées (théorie des cognèmes, interaction syntaxique distribuée, Bottineau 2002, 2005, 2007b) ou des phases et dynamiques de l'activité de la production de la parole qui marquent d'ailleurs l'inscription corporelle du langage (Bottineau 2007a). Un contrat blanc de l'ANR *Perception sémiotique et socialité du sens : formes, sens commun, activités symboliques* vient d'ailleurs en support de certaines de ces orientations (ANR-06-BLAN-0281).

Nous proposons de réunir dans une même approche construction de théories et modèles affine aux linguistiques cognitives et aux courants dynamiques et 'culturalisants' des sciences cognitives et travail expérimental. Un ensemble non exhaustif d'actions déjà en cours ou en phase avancée de définition figure ci-dessous.

## 2.a. Actions

### 1) Théorie physiognomique du geste interne

- ***Dynamique concrète des verbes de mouvement***, ses corrélats sensorimoteurs et électrophysiologiques (neurones miroirs & embodiment) ; procédure : un mot est affiché à l'écran, puis s'efface, et le sujet doit indiquer avec son doigt l'emplacement de l'affichage ☞ en utilisant l'écran tactile.
- ***Geste co-verbal*** : évaluer l'hypothèse que la symbolisation linguistique et la symbolisation gestuelle prennent leur source dans le même espace physiognomique.
- ***Évaluer l'hypothèse de l'imbrication des systèmes verbal et gestuel*** en associant production verbale et gestuelle dans le contexte d'une tâche de dénomination d'objets conduite avec des patients présentant une grande difficulté à nommer des objets dont ils sont souvent capables de montrer l'usage (anomie aphasique). ☞ Importance pour la rééducation.

## 2) Perception sémiotique et socialité du sens : formes, sens commun, activités symboliques [Projet blanc ANR-06-BLAN-0281]

Ce projet à caractère fortement *interdisciplinaire* propose de nouvelles voies pour faire avancer ensemble sciences cognitives, sciences du langage et sciences sociales. La clé en est de rendre compatibles deux caractéristiques fondamentales du sens : sa *perceptibilité* et sa *socialité*, et de les faire opérer dans le cadre d'activités *symboliques*. Plus précisément : (i) la socialité du sens doit être rapportée d'emblée à des formes et des activités symboliques, qui redirigent en permanence les interactions et conditionnent la formation des valeurs et des utilités, (ii) le sens en tant que social ne se sépare pas d'une recherche d'expression, concomitante de la formation de divers médias sémiotiques et d'une constante ritualisation des conduites, fondant la possibilité de la répétition et d'une évaluation des écarts, (iii) l'historicité et la socialité du sens trouvent leur répondant, au niveau de l'expérience individuelle, dans une perception d'emblée *sémiotique*, qui ne se sépare pas de dispositions expressives étroitement dépendantes des médiations sémiotiques instituées.

## 3) L'inscription corporelle du langage social

Dans le cadre d'une théorie de l'acte de langage comme procédure distribuée de co-construction du sens (constitution par la parole de l'idée pour soi et autrui), prise en compte des paramètres physiologiques et environnementaux où s'inscrit l'émergence langagière aux différentes échelles historique (émergence / évolution, acquisition et discours) et étude de la formation de systèmes individués semi-autonomes (démotivation / remotivation bouclée, autopoïèse) :

- **Langage et respiration** : rôle de l'alternance respiratoire dans la discontinuité énonciative (structure et limites des énoncés, syntagmes, lexèmes) et du dialogisme (réversibilité, planification dialogique inscrite dans la morphosyntaxe et la prosodie). Langage, respiration, émotion : affect respiratoire des émotions, effets sur la prosodie, rôle de la prosodie comme compensateur de la perturbation émotionnelle et comme catalyseur et ajusteur (émotionnel et dialogique) de la planification énonciative. Polyfonctionnalité et ubiquité des facteurs (prosodie comme contrôleur respiratoire, comme interface émotion / énonciation, comme segmenteur syntaxique ou lexical (tonèmes) et planificateur dialogique, comme fait externe et interne au langage).
- **Langage, articulation, audition** (proprioception et perception environnementale) **et multimodalité** (gestualité visuelle de l'articulation et autres kinésies) : systèmes phonologiques (Diver 1979, Tobin 1997 : *Phonology as Human Behaviour*, the Columbia School of Linguistics) : propriétés sensori-motrices des systèmes dans les langues et modélisation des articulations sémantiques dérivées (théorie des cognèmes) ; phonosymbolisme et démotivation dans le lexique et la morphologie (idéophones et classificateurs, éléments formateurs des grammèmes) ; morphologie de catégorisation du lexique en classes d'expérience par propriétés enactives et cadrage de la mise en œuvre discursive des lexèmes (polysémie, métaphore, couplages harmoniques ou transgressifs, langue des jeunes)

## 4) Syntaxe et morphologie en typologie linguistique

- **Étude de l'ordre des mots et syntagmes pris pour indices** 1) de l'ordre de construction sémantique dans la procédure d'idéation et 2) de la distribution des contrôles et cibles des procédures d'orientation. Résultats obtenus : trois « types cognitifs » ont été déterminés, a) opérande / opérateur (anglais et langues romanes pour l'essentiel), b) focale / pivot verbal / arguments (breton) et analyse / synthèse (basque). Réduction de

l'écart gotique / celtique > allemand / breton (convergence typologique). Le japonais (type c) malgré une diachronie hybride, très comparable au basque), plusieurs langues du groupe malayo-polynésien (proche du type b), mais avec regroupement de la focale et du pivot verbal en un constituant unique, et polycatégorialité des unités lexicales) et du groupe celtique (proche de b) également, mais sans la polycatégorialité dans le lexique). A long terme, projet de décrypter le schème syntaxique de l'inuktitut.

- **Les grammèmes** affixes, éléments formateurs agglutinables, mono- ou polyfonctionnels, formes sagittales (Hagège 1982) comme marqueurs d'une procédure d'orientation déterminée et, selon les cas et types, de contrôleur et/ou cible spécifiés (distributionnalité) ; inscription des théories de l'invariance (Guillaume 1964, Culioli 1973, Adamczewski & Delmas 1982) et du reprofilage (Cadiot & Visetti 2001) dans le cadre d'une dynamique distribuée. Structure submorphologique des opérateurs (théorie des cognèmes) dans ses rapports à l'expérience sensori-motrice. Inscription en cours de la cognématique dans les courants de la corporéité cognitive et du langage distribué. Remise en cause de la compositionnalité morphologique et mise en évidence de réseaux dont les parties sont manifestées par des marqueurs symbolisant la variation configurationnelle de l'ensemble (basque).
- **Couplage morphologie / syntaxe** : Détermination du statut cognitif en termes d'orientation de faits morphosyntaxiques saillants (accords progressifs et régressifs, harmonies, rôle syntaxique des cognèmes, etc.).

## 5) Parole et diachronie

- La notion de « grammaire émergente », développée dans la mouvance cognitive, remet au premier plan l'acte de parole par rapport au système abstrait, selon une conception qui est déjà celle de W. von Humboldt : « La langue, considérée dans son essence réelle, est quelque chose de passager, continuellement et à tout moment. [...] Elle-même n'est pas une œuvre (érçon) mais une activité (énérgéia). »
- On voudrait donc confronter et faire se confronter les méthodes en diachronie, tout en interrogeant les (r)apports de l'empirisme et du théorique, de la philologie et de la linguistique. On défendra que la théorisation du changement se fait mieux si l'on considère le langage comme activité (énérgéia), conjoignant universaux cognitifs et marqueurs de différenciation sociale et individuelle.

## 6) Émergence du langage, émergence et développement des langues

- Polyfonctionnalité originelle, constitution d'un paradigme d'actions symboliques coordonnées par détournement et fédération d'activités diverses et non reliées ; langage et communautés, cultures, identités ; individuation et autopoïèse.
- Étudier l'hypothèse d'une relation de co-constitution / co-évolution entre le monde visuel chez l'homme et le langage en s'appuyant sur : (i) le statut et qualité à part du monde visuel chez l'homme, (ii) le statut du champ visuel comme medium mnémotechnique, (iii) la possibilité d'une corrélation « phylogénétique » et/ou de détermination réciproque entre vision et langage chez l'homme, (iv) caractère essentiellement visuo-spatial de l'imaginaire linguistique, (v) forte dépendance lexicale - discrimination perceptive.

## **2.b. Résultats attendus**

- Dossier spécial « Synesthésie, langage, métaphore » dans la revue *Intellectica* pour 2009/2010. Ouvrage général sur les synesthésies en 2010.
- *Théorie physionomique du geste interne* : première vague des résultats expérimentaux sur la dynamique concrète des verbes de mouvement en 2009, bilan et publications 2010 ; deuxième vague d'expériences 2010/2011. Première vague des résultats des travaux sur le geste co-verbal 2009/2010. Bilan et publication 2010/2011.
- Colloque international clôturant les travaux du programme blanc de l'ANR [ANR-06-BLAN-0281] *Perception sémiotique et socialité du sens : formes, sens commun, activités symboliques* fin 2009. Ouvrage collectif en 2010.
- Organisation d'un colloque international *Langage et monde visuel*, en 2011.
- Une thèse sur la théorie physionomique du geste interne est programmée dans les mois qui viennent.
- Constitution d'un questionnaire unifié appliqué au fait langagier, son fonctionnement individuel et social, son émergence et évolution individuelle, communautaire et historique, sa multifonctionnalité (rôle dans la formation de la culture, la subjectalisation) ; coordination des modèles théoriques pertinents (cognitifs, distribués) et de la typologie linguistique. Inscription des propositions personnelles (dont les cognèmes) dans un paradigme collectif fédérant plusieurs courants actuels convergents. Les collaborations requises sont en cours d'évolution rapide (à travers la participation à l'école d'été « enaction », aux réunions du DLG – et publications qui en découlent. De ces associations vont ressortir les projets matériels pour l'horizon 2009-2012 (ouvrage collectif et colloque).



### 3.3.2. COMPOSANTE *DYNAMIQUES*

#### **Coordination Christian Hudelot & Françoise Gadet**

On opposait naguère deux façons de faire de la linguistique, modélisation et formalisation vs exploration du savoir et du fonctionnement d'un locuteur. La notion de corpus a largement modifié une telle vision des choses, mais il reste quelque chose de l'opposition dans l'existence d'une perspective tournée vers l'agent usager du langage et des langues, ses savoirs, ses pratiques, ses stratégies, ses modes d'acquisition et de fonctionnement (ou dysfonctionnement) dans les usages ordinaires ou plus exceptionnels.

C'est dans une telle perspective que s'inscrit le secteur Dynamiques, qui traite des liens que le locuteur entretient avec la langue, dans ses relations avec le monde et les propos d'autrui dans la diversité des situations interlocutives. Les corpus que nous constituons, qui permettent de documenter ces acquisitions et ces usages, sont crucialement situés, produits par des agents parlants bien particuliers, interagissant avec d'autres agents, dans des situations identifiables. Et donc leurs conditions de recueil font autant que leur exploitation l'objet d'une grande attention.

Le secteur Dynamiques insiste sur l'hétérogénéité constitutive des pratiques langagières, liées d'abord au fait que les locuteurs ne vont pas de la langue à une mise en mots conforme à des usages, mais de situations discursives à l'appropriation de différentes sortes de mises en mots. Hétérogénéité due également au fait que la langue est au mieux une construction stabilisée du linguiste, quand le locuteur/auditeur concret est confronté à des pratiques discursives dont les différents niveaux, tant au plan formel qu'aux plans conceptuel et fonctionnel, sont soumis à la double pression de l'évolution (synchronie dynamique) et de l'identification ou des normes. Les travaux portent autant sur les processus et la structure des évolutions (grammaticalisation, lexicalisation, complexifications acquisitionnelles, divergence et convergence) que sur les facteurs des dynamiques linguistiques en jeu (zones de fragilité interne, poussées externes des contraintes communicatives, contact avec des langues ou des variétés autres). En effet, le secteur insiste sur le fait que les langues, pas plus que les variétés, les pratiques langagières, les savoirs ou les cultures, ne sont étanches les unes aux autres et que les contacts sont l'un des moteurs des dynamiques linguistique et langagière.

C'est pourquoi on présentera le secteur Dynamiques des langues dans ses deux composantes complémentaires que sont les dynamiques des acquisitions et les dynamiques des usages (plus proches de la psychologie pour la première composante, de la sociologie ou de l'ethnologie pour la seconde), même si la porosité de la frontière se lit à la circulation de chercheurs entre les deux équipes, de même qu'à l'existence de projets inter-équipes. On ne s'étonnera pas alors de trouver, à côté d'opérations de recherche propres à chacune, des opérations communes aux deux équipes. Enfin, dans la mise en relations entre la langue comme appropriation des locuteurs et l'objet qu'ils se construisent et peuvent dès lors manipuler, comme d'autres formes de savoirs, il était important de mettre en relations, dans des opérations trans-secteurs, les dynamiques langagières et les modèles linguistiques et/ou discursifs.

### 3.3.2.1. Equipe *Dynamiques des acquisitions*

**Coordination : Christian Hudelot & Colette Noyau**

#### **Membres de l'équipe *Dynamiques des acquisitions***

<b>Membres rattachés à titre principal</b>	Chercheurs et enseignants-chercheurs	L. Danon-Boileau, C. Hudelot, M. Kihlstedt, M. Leroy, C. Le Cunff, C. Martinot, C. Noyau, C. Parisse, M.-C. Poudet, C. Préneron, C. Romero, E. Veneziano	12	14
	ITA	M. Kugler-Lambert	1	
	Membres émérites ou honoraires	C. Chevrier-Muller	1	
<b>Doctorants</b>	L. Allal, N. Al Huda Al Awbali, A. H. Bah, O. Binganga, K. Bouchez, N. Boukhelouf, F. Caetano-Rousselot, H. Colin, G. Florigny, H. Kourouka-Mitsika, A. Pégaz-Paquet, J. Tan, F. Zirpoli, L. Toti, A. Xenaki		15	
<b>Jeunes chercheurs associés</b>			0	
<b>Membres associés</b>	A. Bensalah-Poulain (orthophoniste), L. Caid (Educ. Nat, La Réunion), J. N. Choi (Univ. Autonome de Barcelone), S. Gerolimich (Trieste), M. Leclere-Messebel (Paris 3), Ch. Maillard (Liège), A. Morgenstern (ENS-LSH), U. Paprocka-Piotrowska (Univ. Catholique de Lublin, Pologne), G. Sanz Espinar (Univ. Autonome de Barcelone)		9	
<b>Collaborations extérieures</b>	M. Berges-Bounes, C. Bernardeau, M.-C. Devaux, C. Ferron, G. Lebugle, E. Lenoble, L. Monnier, A.-M. Pecarelo (tous à l'UPPEA - Hôpital Sainte-Anne, Paris), G. Bernini (Bergame), V. Boiron (IUFM Bordeaux), F. Bourdoux (Liège), T. Bosnjak (Zagreb), C. Checri-Bailly (Ecole d'orthophonie Pitié-Salpêtrière, Paris), G. Chevalier, E. Clark (Stanford), F. Cuisinier (Paris 10), K. Duvignau (IUFM de Toulouse), L. Erbolato Melo (USP São Paulo, Brésil), D. Gayet (Paris 10), R. Hassan R. (Lille 3), B. Héberlé-Dulouard (CASNAV), D.-F. Idiata (Libreville), A. Katsura (Texas Tech University), D. Kugler (traductrice), J. Kuvac (Zagreb), A. Lainé (Centre de Guidance Infanto-Juvenile Toulouse), K. Lindner (Munich), S. Lopez Ornat (Madrid), P. Machado Goyano Mac-Kay (PUC São Paulo), S. C. Marquesi (PUC São Paulo), E. Mathiot (Lille 3), C. Maury Rouan (Provence-Aix-Marseille), L. McCune (Rutgers), M.-A. Morel (Paris 3), E. Nicolescu (Spiru Haret), J. Perissinoto (UNIFESP São Paulo), A. Peters (Hawaii), A. Philippe (INSERM, Paris 5), C. Rojas (UAM Mexico), N. Salagnac (IUFM de Gravelines), M. Touzin (Hôpital du Kremlin Bicêtre), A. Viberg (Uppsala), R. Vion (Provence-Aix-Marseille), E. Zaretsky (Massachusetts)		44	

#### **Présentation générale de la problématique scientifique**

Les études sur l'acquisition linguistique connu un essor important dans les 20 dernières années, pour l'acquisition des langues premières comme pour celle des langues secondes ou étrangères et l'acquisition bilingue du langage. Ces travaux ont bénéficié d'une approche fonctionnaliste et émergentiste. Les catégories ne sont pas données comme préexistantes, mais comme à construire par l'enfant ou l'apprenant dans et pour ses interactions avec des locuteurs et avec la langue dans ses activités de compréhension et de production en contexte. On est ainsi amené à prendre comme point de départ la dimension textuelle ou discursive au sein de laquelle les formes prennent fonction et sens.

Les dynamiques qui se jouent dans les situations d'utilisation de la langue en cours d'appropriation relèvent et de la construction du système linguistique, et de l'acquisition des modes d'usage de la langue pour produire et pour comprendre, dans des contextes et des tâches déterminés.

La complémentarité entre linguistes, psycholinguistes, chercheurs de terrain associés aux structures de la recherche et de l'intervention thérapeutique, et théoriciens des applications didactiques permet de parcourir dans sa totalité la chaîne des observables aux interventions.

Le fort ancrage international des recherches en acquisition de L2 et du bilinguisme, tant en Europe (Réseau européen CASLA, Espagne, Pologne, Suède), qu'ailleurs (Canada et Afrique/Océan Indien francophones notamment). Ce qui permet, au plan des comparaisons typologiques, de compter avec un vaste éventail de langues impliquées et d'expertise linguistique sur les L1 et les L2, et au plan des situations linguistiques considérées, de contraster les contextes sociolinguistiques et culturels d'acquisition.

### **Liste des opérations**

3.3.2.1.1. *Acquisition initiale du langage*

3.3.2.1.2. *Acquisition des langues secondes ou étrangères et du plurilinguisme*

3.3.2.1.3. *Pathologies de la communication et du langage et remédiations*

### **3.3.2.1.1. Acquisition initiale du langage**

#### **1. Membres de l'opération**

*Membres rattachés à titre principal (10 : 2 PR, 5 MC, 1 DR, 1 CR, 1 IE)*

Christian Hudelot (responsable)

Edy Veneziano (responsable)

Marie Carcassonne

Claude Chevrie-Muller

Mireille Froment

Marie Kugler-Lambert

Marie Leroy

Claire Martinot

Christophe Parisse

Clara Romero

*Doctorants (5)*

Nour Al Huda Al Awbali

Abdul Hamidy Bah

Naziha Boukhelouf

Hélène Colin

Hélène Kourouka-Mitsika

Anne Pégaz-Paquet

*Membres associés (4)*

Sonia Gerolimich (Université de Trieste)

Malory Leclere-Messebel (Université Paris 3)

Aliyah Morgenstern (ENS-LSH)

Urzsula Paprocka-Piotrowska (Université catholique de Lublin)

*Collaborations extérieures (18)*

Tomislava Bosnjak (Université de Zagreb)

Eve Clark (Université de Stanford)

Lélia Erbolato Melo (USP São Paulo)

Rouba Hassan (Université de Lille 3)

Aoyama Katsura (Texas Tech University)

Jelena Kuvac (Université de Zagreb)

Katrin Lindner (Université de Munich)

Susana Lopez Ornat (Université de Madrid)

Ana Paula Machado Goyano Mac-Kay (PUC São Paulo)

Sueli Cristina Marquesi (PUC São Paulo)

Emmanuelle Mathiot (Université Lille 3)

Lorraine McCune (Rutgers University)

Elisabeta Nicolescu (Université Spiru Haret de Bucarest)

Jacy Perissinoto (UNIFESP São Paulo)

Ann Peters (Université d'Hawaii)

Cecilia Rojas (UAM Mexico)

Nathalie Salagnac (IUFM Gravelines)

Elena Zaretsky (Université du Massachusetts, Amherst)

## **2.a. Actions**

### **1) Emergence de la Grammaticalité chez le jeune enfant : Facteurs cognitifs, linguistiques et conversationnels**

Lorsqu'ils commencent à parler, les enfants produisent tout d'abord un mot à la fois. Or, autour de l'âge de trois ans, ils semblent avoir à leur disposition les outils nécessaires à la production d'énoncés relativement complexes et grammaticalement appropriés. L'émergence de la grammaticalité est un élément crucial pour comprendre la nature de cette évolution. Comment expliquer que les énoncés des enfants commencent à contenir des morphèmes reconnaissables tels que les déterminants, les pronoms et les auxiliaires, que les formes verbales soient employées de manière contrastive, que noms et verbes développent des patterns d'occurrence distinctifs, et que les énoncés contiennent majoritairement plusieurs mots ?

L'étude de l'émergence de la grammaticalité a été abordée de diverses manières. Selon les approches générativistes, les enfants sont supposés s'appuyer sur des savoirs innés, contraints par des principes linguistiques, qui peuvent être généralisés. Selon les approches constructivistes, le langage est construit par petits bouts à partir de l'élaboration cognitive du langage que l'enfant s'approprie et avec lequel il interagit. Ce projet a pour objectif de contribuer à ce débat grâce à une étude novatrice et originale de l'émergence de la grammaticalité. Il s'agit d'allier le recueil et l'analyse détaillée de données longitudinales en milieu naturel à des méthodes expérimentales, ce que l'on ne trouve pas dans les travaux précédents. La complémentarité de ces approches permettra d'étudier à la fois la production et la compréhension des enfants, d'analyser les caractéristiques pertinentes du langage adressé aux enfants ainsi que les échanges conversationnels, et de tester, grâce à une méthode d'entraînement, une hypothèse selon laquelle les énoncés à plusieurs mots sont favorisés par l'enchaînement dialogique d'énoncés à un mot.

De plus, cette recherche pourra bénéficier de la collaboration d'experts de pointe dans les domaines étudiés. Ces derniers contribueront et travailleront sur des données en anglais et en espagnol. Cette approche multilingue ouvrira les horizons du projet et permettra de comparer les résultats obtenus pour le français avec ceux d'autres langues, les intégrant sur le plan de la recherche internationale.

Le projet comprend :

1. Un outil évaluatif des premières acquisitions grammaticales grâce à l'analyse de trois phénomènes : (a) les « fillers » dont on pense qu'ils permettent d'initier un processus menant à la morphologie libre ; (b) la production des verbes avant et après l'apparition de formes morphologiquement différentes pour les mêmes verbes types ; (c) les débuts et l'évolution des énoncés à plusieurs mots.

2. Une étude détaillée des processus et mécanismes sur lesquels les enfants s'appuient pour construire leur grammaire, grâce à l'étude des morphèmes pré-nominaux et pré-verbaux ainsi que des formes verbales attestées dans le langage adressé à l'enfant et dans la conversation mère-enfant. Des types particuliers d'échanges conversationnels que l'on considère comme centraux dans la transition vers la production d'énoncés à plusieurs mots seront analysés dans les données longitudinales et seront mis en œuvre à travers une procédure expérimentale afin de tester spécifiquement cette hypothèse.

Les analyses seront menées sur des données longitudinales provenant d'enfants en train d'acquérir le français, l'anglais et l'espagnol. Un entraînement expérimental sera mis en place sur 20 enfants « expérimentaux » et 10 « contrôles » pour chaque langue.

Nous prévoyons que les analyses simultanées des fillers et des formes verbales montreront que la différenciation entre noms et verbes se produit quand les énoncés à plusieurs mots deviennent le mode d'expression dominant chez l'enfant. C'est aussi à ce moment-là que de nombreux verbes seront utilisés avec des formes morphologiquement contrastées. Nous stipulons aussi que les relations entre les caractéristiques du LAE et de l'activité conversationnelle d'une part, et les caractéristiques des fillers et des formes verbales de l'autre, ne sont pas aléatoires. En conformité avec les hypothèses, nous nous attendons à ce que les échanges conversationnels encouragés dans les séances d'entraînement permettent d'acquérir la production d'énoncés à plusieurs mots. Cette recherche pourrait dès lors avoir des retombées directes sur les méthodes utilisées en remédiation du langage.

Nous pensons par ailleurs trouver des relations complexes entre les différents aspects étudiés confrontant ainsi le point de vue constructiviste sur l'acquisition du langage, selon lequel la signification d'éléments acquis à un moment donné peut être modifiée à la lumière d'éléments acquis par la suite. Nous pensons que la comparaison des résultats obtenus sur le français, l'anglais et l'espagnol contribuera à une meilleure compréhension de l'émergence de la grammaticalité ainsi qu'aux connaissances multilingues.

## **2) Acquisition et reformulation**

Ce projet, déjà en cours et qui a fait l'objet de deux publications et de plusieurs présentations et conférences, a pour objectif de rendre compte du mode de construction de la prédication au cours de la période dite des acquisitions tardives (4-10 ans), période peu documentée du point de vue de l'acquisition de la langue orale.

L'hypothèse défendue ici est que l'acquisition de la langue est déterminée par des procédures de reformulation que l'enfant, en fonction de ses compétences langagières, applique à un énoncé entendu et mémorisé, pour produire un autre énoncé, aussi bien de sens équivalent que différent, mais qui partagera avec le premier énoncé un aspect invariant (au niveau du lexique, de la construction, ou du sens) et qui comportera en même temps un changement concernant

un autre niveau que celui qui est invariant. La description et l'analyse de l'ensemble des phénomènes liés à la construction du sens ou de la prédication se font par comparaison entre un texte source et les productions des enfants qui racontent à leur tour le même texte (fiction de 500 mots environ). Cette hypothèse d'une acquisition par reformulation/transformation, s'appuie sur le postulat que les enfants se livrent, de cette façon, à une investigation du fonctionnement de leur langue qui correspond à des mécanismes communs à toutes les langues (quelles contraintes déterminent la co-occurrence de catégories de mots, comment un mot d'une catégorie sélectionne un mot plutôt qu'un autre pour produire une signification, comment augmente la complexité des phrases élémentaires par réduction et grammaticalisation), et à des mécanismes spécifiques à chaque langue.

**Volet 1** : Les stades d'acquisition dans différentes langues maternelles.

Les mécanismes communs ou spécifiques aux langues sont recherchés dans la (re)construction faite par les enfants des prédications simples (le verbe et ses arguments) du texte source, des prédications liées (prédication simple suivie d'une relative) et des prédications complexes (prédication principale et seconde dans la même phrase). L'hypothèse est que les 15 enfants de chaque groupe d'âge (4, 6, 8 et 10 ans) reformulent les séquences correspondantes du texte source de façon différente et que cette façon nous renseigne sur ce qu'ils savent faire avec la langue pour produire un sens donné. L'expérimentation dans 9 langues différentes (français, italien, roumain, allemand, anglais, croate, polonais, grec, arabe) devrait montrer dans quelle mesure l'hypothèse acquisitionnelle de la reformulation est généralisable.

**Volet 2** : Constitution d'un ensemble de normes de références : Observatoire du français oral chez les enfants scolarisés en France.

Le corpus (pour la partie française) a été recueilli auprès de 60 enfants français ou francophones natifs de 4 à 10 ans et constitue donc un ensemble de données qui pourrait servir de référence aux acteurs scolaires impliqués dans l'enseignement du français langue maternelle. Il permet en particulier de voir à quel âge une compétence particulière commence à être attestée, ce qui permet que l'enseignement s'appuie sur cette compétence en construction.

**Volet 3** : Reformulation et dysphasie.

La troisième étape est d'utiliser un protocole expérimental comparable à celui décrit ci-dessus en l'adaptant aux enfants dysphasiques qui ne semblent pas disposer des mêmes compétences de reformulation que les enfants sans troubles de langage. Cependant dans une perspective de remédiation, la consolidation de certaines procédures de reformulation pourrait être envisagée en plus des outils déjà disponibles qu'utilisent les orthophonistes.

### **3) NarraToM : Narration et Théorie de l'esprit**

Le but principal du présent projet est d'évaluer de différentes manières les compétences des enfants en matière de langage et de théorie de l'esprit, ainsi que d'étudier le rôle de différentes méthodes d'intervention sur l'amélioration des performances des enfants, qu'il s'agisse d'acquérir des nouvelles compétences ou de rendre accessibles celles déjà existantes.

Pour ce faire, nous avons choisi d'investiguer une compétence cognitivo-langagière très souvent sollicitée dans la vie quotidienne et qui, par son aspect de décontextualisation, est centrale aussi dans la réussite scolaire, celle de raconter et de produire des récits.

L'originalité de cette recherche est d'essayer d'apporter un nouvel éclairage à ces questions, par deux moyens principaux :

a) d'une part nous mettrons en place un système d'évaluation multiple où les mêmes compétences sont sollicitées dans des conditions diverses d'accessibilité. Ainsi l'état de

connaissance de l'enfant sera évalué: sans soutien externe (premier récit spontané) et après que l'enfant a été exposé à différents types de soutiens externes dont certains l'impliquent directement en tant que partenaire d'interaction. Il sera alors possible de mieux déterminer si le manque de compétence doit être attribué à des variables développementales ou s'il peut être reconsidéré, même de manière radicale, comme due à des variables d'une autre nature;

b) d'autre part, nous comparerons différentes méthodes d'intervention dont l'efficacité sera évaluée en termes d'amélioration de l'une ou l'autre des différentes composantes prises en compte dans l'analyse, autant sur le court terme qu'à plus longue échéance. Comparées elles-mêmes à une situation contrôle où les enfants sont amenés à re-raconter l'histoire après une familiarisation supplémentaire avec les images, nous pourrions évaluer le rôle des méthodes d'interventions par rapport au fait de raconter une nouvelle fois l'histoire.

Dans l'esprit de mieux cerner l'état des compétences et la marge possible de progression, les enfants à développement typique seront comparés à des enfants présentant des troubles du langage (enfants dysphasiques) mais aussi à des enfants présentant des troubles de la communication et/ou du développement cognitif (enfants autistiques et trisomiques).

La même étude effectuée chez des enfants parlant des langues différentes pourra nous aider à comprendre dans quelle mesure la non-expression d'aspects évaluatifs dans une langue donnée peut être due à des difficultés morphosyntaxiques spécifiques à cette langue, et inversement si l'absence de telles difficultés dans une autre langue amènerait l'enfant à les exprimer plus précocement.

La recherche dans son ensemble fournira une base de données permettant la mise en place d'un outil d'évaluation nuancée sur laquelle baser ultérieurement des épreuves standardisées.

#### **4) Développement grammatical de l'enfant de 1 à 5 ans**

Le développement grammatical de l'enfant fait ici référence à la période qui va de l'âge où les enfants produisent de simples mots isolés (de 1 à 2 ans selon les enfants) à celui où ils produisent des énoncés qui sont très proches de ceux des adultes qui les entourent (pour cette recherche, on se limite aux âges préscolaires, jusqu'environ 5 ans).

Cette période présente plusieurs moments cruciaux, en particulier d'une part, le début de l'assemblage de mots et de production de marques grammaticales (autour de 18 à 24 mois), et d'autre part l'augmentation importante de la qualité du langage (de 2 à 3-4 ans) où les enfants accèdent à une grande richesse syntaxique et à un langage de qualité (notamment par la qualité de prononciation, la longueur et le volume des énoncés).

Cette action est soutenue par deux ANR (Emergram 2006, Polycat 2006) dont le laboratoire *MoDyCo* est partenaire principal ou secondaire. Elle est aussi impliquée dans le projet Leonard. Elle est remarquable par le nombre de personnes qui y sont impliqués (dans et hors du laboratoire). Elle comporte plusieurs volets qui se complètent :

*Recueil original de corpus longitudinaux sur une longue période allant au moins de 18 mois à 3 ans.* Ce recueil et l'important travail de transcription qu'il suppose ne peut se mettre en place que grâce aux contrats ANR obtenus et à la mutualisation des moyens offerts par tous les projets financés. Toutes les transcriptions suivent le format CHILDES et sont accompagnées de document vidéo.

*Emergence de la grammaire.* Autour de l'âge de 2 ans, les enfants commencent en même temps à combiner des mots isolés dans un même énoncé, à produire des marques syntaxiques attestant d'une certaine maîtrise de catégories grammaticales comme le nom, le verbe, et à

produire des enchaînements conversationnels réussis. L'hypothèse est que la simultanéité de ces manifestations n'est pas le fruit du hasard mais qu'elle se retrouve chez tous les enfants. De plus, on cherchera à vérifier de manière expérimentale si les prémisses de ces évolutions peuvent être identifiées chez les enfants dès le début de leurs productions complexes. Ceci n'est possible que grâce au recueil longitudinal mis en place dans cette action.

*Développement des marqueurs syntaxiques.* Après l'âge de 2 ans, la grammaire des enfants continue de progresser à grands pas et de nombreux indicateurs permettent de suivre ce développement. Une recherche porte en particulier sur les prépositions qui sont typiques de cette période du développement. Cette période est également très importante car elle permet de contrôler l'apparition d'éventuels troubles de développement du langage, en comparant les enfants entre eux.

### **3.3.2.1.2. Acquisition des langues secondes ou étrangères et du plurilinguisme**

#### **1. Membres de l'opération**

*Membres rattachés à titre principal (3 : 1 PR, 2 MC)*

Colette Noyau (responsable)

Maria Kihlstedt

Catherine Le Cunff

*Doctorants (3)*

Guilhem Florigny

Jia Tan

Francesca Zirpoli

*Membres associés (3)*

Jin Nam Choi (Université de Barcelone)

Urszula Paprocka (Université Catholique de Lublin, Pologne)

Gema Sanz Espinar (Université Autonome de Madrid)

*Collaborations extérieures (4)*

Giuliano Bernini (Université de Bergamo)

Gisèle Chevalier (Université de Moncton)

Karine Duvignau (IUFM de Toulouse)

Ake Viberg (Université d'Uppsala)

#### **2. Problématique**

Cette opération, qui regroupe l'ensemble des recherches prévues sur les processus d'acquisition de langues postérieures à l'acquisition primaire du langage, se structure en trois actions, l'une centrée davantage sur les questions théoriques de linguistique de l'acquisition (Acquisition des langues, plurilinguisme, et typologie des langues : approche comparative, coord. C. Noyau), l'autre centrée sur une population particulière, aux confins entre acquisition primaire du langage et acquisition de langues secondes, portant sur les enfants en acquisition précoce du bilinguisme (coord. M. Kihlstedt), et la dernière sur la construction de connaissances scolaires en L1 ou L2, du point de vue de l'interaction entre développement langagier, travail cognitif, et communautés d'apprentissage (coord. C. Le Cunff).



Les recherches prévues dans cette action s'inscrivent en bonne part dans le prolongement du projet « Représentation des événements : sémantique verbale, temporalité et développement des énoncés complexes » du quadriennal précédent. Il s'agit d'approfondir, sur une variété de situations d'acquisition des langues et du bi/plurilinguisme, les modalités de la conceptualisation (le 'penser pour parler', cf. Slobin 1995), en comparant des productions textuelles orales (parfois aussi écrites) pour une même tâche, en L2 entre étapes de l'acquisition, entre locuteurs apprenants de groupes homogènes, entre la L2 et la L1, en L2 pour des locuteurs de différentes L1. Il s'agit ainsi de cerner comment la L1 et les L2 acquises précédemment infléchissent les processus d'acquisition, et comment la L1 ou la langue dominante chez l'individu bilingue infléchit les modes de conceptualisation pour la mise en mots.

## **2.a. Actions**

### **1) Acquisition des langues, plurilinguisme, et typologie des langues : approche comparative**

Les recherches prévues dans cette action s'inscrivent en bonne part dans le prolongement du projet « Représentation des événements : sémantique verbale, temporalité et développement des énoncés complexes » du quadriennal précédent. Il s'agit d'approfondir, sur une variété de situations d'acquisition des langues et du bi/plurilinguisme, les modalités de la conceptualisation (le 'penser pour parler', cf. Slobin 1995), en comparant des productions textuelles orales (parfois aussi écrites) pour une même tâche, en L2 entre étapes de l'acquisition, entre locuteurs apprenants de groupes homogènes, entre la L2 et la L1, en L2 pour des locuteurs de différentes L1. Il s'agit ainsi de cerner comment la L1 et les L2 acquises précédemment infléchissent les processus d'acquisition, et comment la L1 ou la langue dominante chez l'individu bilingue infléchit les modes de conceptualisation pour la mise en mots.

Les domaines privilégiés de cette action sont :

- **La sémantique verbale** : développement et structuration du lexique verbal pour la représentation des procès. Ce volet regroupe des recherches sur la représentation des procès en L2 vs L1 ou en La vs Lb à partir de productions langagières calibrées (récits sur images ou support vidéo), en collaboration avec le Réseau Européen 'Comparative Approach of Second language Acquisition (CASLA) et le Pr. A. Viberg (U. d'Uppsala), et une recherche expérimentale en coopération avec le laboratoire J. Lordat de Toulouse (ACI-JC dir. par Karine Duvignau) sur : *Architecture structurale et fonctionnelle du lexique verbal : La flexibilité sémantique comme principe fondamental de la cognition humaine et artificielle*, portant sur des populations francophones d'enfants sans et avec troubles du langage, d'adultes sans et avec pathologies, et sur des sujets en français L2 et de diverses L1 (l'équipe de Nanterre effectue l'étude sur des locuteurs enfants et adultes coréanophones en L1 et en L2, et des locuteurs enfants et adultes suédophones en L1 et L2).

- **Les énoncés complexes** : développement des formes syntaxiques et des fonctions sémantiques, discursives et pragmatiques de la parataxe et de l'hypotaxe en L2 (ou en La et en Lb chez les bilingues).

Une étude portera également sur la temporalité (morphologie temporelle et structure narrative) en L1 suédois / français, 2 L1 suédois-français et L2 français chez des enfants suédophones.

## Méthodes

Toutes les études menées dans le cadre de ce projet (thèses en cours) ont recours à la méthode comparative, à partir de productions textuelles analysées selon une optique fonctionnaliste.

Une étude spécifique menée avec G. Chevalier (U. de Moncton, N.-Br., Canada) portera sur le niveau avancé du français L1 minoritaire en environnement anglophone) : étude sur la diversification du lexique verbal chez les jeunes Acadiens entrant à l'université et devant s'approprier une variété de français académique adéquate pour la construction des connaissances scientifiques et techniques et pour la maîtrise communicative exigée par leur devenir professionnel de cadres.

En face des données suscitées, on examine comment les domaines linguistiques en jeu sont traités ou actifs dans des interactions de classe de diverses matières scolaires, et interviennent dans la construction des connaissances disciplinaires parallèlement au travail d'appropriation de la langue thématisé ou non.

## 2) Acquisition précoce de L2 ou du bilinguisme

Projet sur le bilinguisme au Lycée Français de Stockholm en cours de montage, en collaboration avec Suzanne Schlyter et Jonas Granfeldt à l'université de Lund, Suède. Cette action s'inscrit dans le réseau franco-suédois géré par *MoDyCo* depuis 2005. Le projet associe, par sa problématique, l'acquisition de L1 et de L2.

**Population** : 14 enfants âgés de 4 à 7 ans se répartissent sur les trois groupes suivants (à compléter pendant les années à venir)

- Français langue maternelle (1 L1)
- Français et suédois langues maternelles depuis la naissance (2 L1)
- Suédois langue maternelle, acquisition du français à partir de 4 ou 6 ans (enfant L2)

La problématique est celle de la relation entre l'acquisition d'une L1, de 2 L1 et d'une L2 à un âge précoce. Il est généralement admis qu'il existe des différences nettes entre l'acquisition du français L1 et L2 : en L1, elle se déroule typiquement sans erreurs, entraves et retours en arrière, contrairement à ce qui est typiquement le cas en L2 et notamment chez les adultes. Ces différences ont été attestées en particulier pour les phénomènes suivants :

- genre (attribution et accord)
- les pronoms clitiques (forme et position)
- accord sujet-verbe
- temps et aspect.

En revanche, les résultats sont moins concluants en ce qui concerne les enfants bilingues (simultanés et successifs). Deux questions se posent initialement :

Un enfant qui acquiert simultanément deux langues se comporte-t-il, dans sa langue non dominante, comme a. un enfant monolingue avec un certain retard (=2 L1) ou b. comme un adulte L2 (= 1 L1 + 1 L2) sur les points mentionnés ci-dessus ?

Un enfant qui démarre l'acquisition d'une L2 à un âge précoce (le bilinguisme successif) se différencie-t-il d'un enfant issu du bilinguisme simultané sur les points mentionnés ?

Les premières observations donnent à penser que les enfants issus d'un bilinguisme simultané ressemblent plus à des enfants monolingues, mais avec un certain retard dans la langue non dominante, alors que les enfants issus du bilinguisme successif manifestent initialement davantage de traits typiquement « L2 ». (cf. Granfeldt, Schlyter, Arapian & Kihlstedt 2007).

On travaillera également sur les phénomènes d'étayage en langue forte ou faible entre adultes et enfants et entre enfants chez ces populations, en liaison avec l'opération Interactions langagières et acquisitions.

### **2.b. Résultats attendus**

On escompte de cette opération des avancées sur le développement acquisitionnel, particulièrement aux stades avancés, et spécifiquement en ce qui concerne le développement du lexique verbal et l'adéquation de la représentation des événements, et le développement des énoncés complexes, ainsi que sur la maîtrise relative de la La et de la Lb en situation de bilinguisme, particulièrement dans ces mêmes domaines du lexique des procès et des énoncés complexes.

Des liens seront établis avec les travaux menés dans le cadre de l'opération *Syntaxe et sémantique grammaticale descriptives*.

Organisation de plusieurs journées d'étude internationales, d'un colloque final, et rédaction d'un volume de synthèse sur le développement de la représentation des événements et la complexification du discours en L2 et chez les sujets apprenants plurilingues.

Participation à la concrétisation des retombées didactiques de nos résultats pour l'élaboration d'outils d'apprentissage institutionnel à différents niveaux, notamment pour l'éducation en français minoritaire en Amérique du nord (en particulier avec certains des partenaires extérieurs de l'opération transversale *Dynamiques acquisitionnelles et sociolinguistiques des situations de la francophonie périphérique*).

### **3.3.2.1.3. Pathologies de la communication et du langage et remédiations**

#### **1. Membres de l'opération**

*Membres rattachés à titre principal (5 : 1 PR, 3 CR, 1 IE)*

Laurent Danon-Boileau (responsable)

Christophe Parisse (responsable)

Marie Kugler-Lambert

Marie-Christine Pouder

Christiane Préneron

*Membre associé (2)*

Yamina Bensalah-Poulain (Orthophoniste)

Christelle Maillart (Université de Liège)

*Collaborations extérieures (20)*

Corinne Bernardeau (UPPEA)

Véronique Boiron (IUFM de Bordeaux)

Françoise Bourdoux (Université de Liège, Belgique)

Marica Berges-Bounes (UPPEA - Unité de Psychopathologie de l'enfant et de l'adolescent de l'hôpital Sainte-Anne à Paris)

Carole Checric-Bailly (Ecole d'orthophonie Pitié-Salpêtrière à Paris)

Marie-Claude Devaux (UPPEA)

Catherine Ferron (UPPEA)

Bénédicte Héberlé-Dulouard (Professeur des Ecoles, Formatrice CASNAV)

Dominique Kugler (Traductrice)

Anne Lainé (orthophoniste, Centre de Guidance Infanto-Juvenile Toulouse)  
G rard Lebugle (UPPEA)  
Evelyne Lenoble (UPPEA)  
Claire Maury Rouan (Universit  de Provence   Aix-Marseille)  
Luce Monnier (UPPEA)  
Mary-Annick Morel (Universit  Paris 3)  
Anne-Marie Pecarelo (UPPEA)  
Nathalie Salagnac (IUFM de Gravelines)  
Monique Touzin (Orthophoniste, H pital du Kremlin Bic tre)  
Anne Philippe (INSERM, Paris 5)  
Robert Vion (Universit  de Provence - Aix-Marseille)

## **2. Probl matique**

L' quipe s'est d finie un ensemble d'objectifs communs :

1. appliquer les m thodes d'observation et d'analyse du dialogue   des corpus issus de la prise en charge d'enfants en difficult  de communication orale et  crite.
2. faire l'inventaire des diff rentes points d'incidence de la pathologie de l'acquisition du langage et des registres o  elle s'organise (dimension pr verbale, communicationnelle et pragmatique, usage des registres syntaxico-s mantiques).
3. rep rer les variations dans l' mergence du trouble linguistique et communicationnel, les conditions dans lesquelles ces difficult s s'estompent/se renforcent.
4. rep rer les indices pr curseurs des troubles de la communication et du langage par le recours   l'analyse de films familiaux d'enfants chez lesquels un trouble av r  a  t  diagnostiqu  ult rieurement.
5. enfin,   la faveur de ces consid rations, d gager les pr -requis de l'acquisition du langage chez l'enfant tout-venant.

De mani re programmatique, nous ferons porter l'essentiel de l'effort sur les enfants pr sentant un trouble de la communication ou du langage. Notre propos sera notamment d'explorer le bien-fond  de diff rentes propositions th oriques pour tenter de comprendre ce qu'il en est de la communication verbale et non verbale chez l'enfant pr sentant des troubles du d veloppement de la communication et du langage.

### **2.a. Actions**

#### **1) Action d pistage pr coce de l'autisme**

Les deux axes d'explication de l'autisme que sont la psychanalyse et la cognition postulent que l'enfant autiste ne dispose pas d'une repr sentation de la repr sentation de l'autre, soit par d ficit soit parce qu'il ne peut pas accepter de se sentir distinct de l'autre.   la lumi re d'observations cliniques, le groupe de travail tente de d finir un autre axe d'explication possible du trouble de la communication non verbale d'un enfant sans langage, qui part de l'id e d'un malaise dans le signe. Communiquer, c'est tenter de faire partager un  tat, une sensation, un affect, un projet, une pens e. Or   l' vidence, on ne peut d signer directement   autrui ce qui est log    l'int rieur de son corps   soi. On peut lui demander de pr ter attention   un bruit,   un  v nement visible, mais on ne peut lui donner acc s directement   une sensation ou une pens e int rieure. D'o  le recours   un objet externe que l'on donne   percevoir en qualit  de signifiant de cette sensation interne. Pour communiquer on d signe

donc à autrui un objet de perception aussi directement audible, palpable ou visible pour lui que pour soi, en faisant le pari que pour lui aussi il vaudra pour l'état intérieur que l'on veut faire passer. Pour prendre un exemple trivial, être triste est un état intérieur que l'on ne parvient à communiquer à autrui qu'en lui montrant quelque chose qu'il peut voir, sentir ou entendre et associer à cet état, larmes ou cris par exemple. Evidemment, pour croire qu'autrui puisse se prêter au jeu et prendre un objet de perception comme signe de sensation chez l'autre, il faut d'abord faire fond pour soi-même sur cette association entre sensation et perception. La capacité d'un sujet à construire des associations de ce type est condition sine qua non à son recours au signe. C'est cette association entre sensation et perception qui fait défaut aux enfants gravement perturbés.

Cela rejoint par exemple ce que visent les cognitivistes lorsqu'ils soulignent le défaut d'intermodalité d'enfants en grande difficulté de communication. Cela reprend aussi la notion psychanalytique de « démantèlement » créée pour désigner cette façon de faire des enfants autistes qui sont capables de manipuler un objet en regardant ailleurs et en écoutant une troisième chose, chacune de ses modalités sensorielles pouvant vivre sa vie de manière indépendante. Et l'on comprend alors que cette tendance l'empêche d'associer sensation et perception. Avec cette conséquence que le signe lui devient impraticable. C'est cette hypothèse que nous tenterons de confronter avec les faits.

Dans cette perspective, l'équipe se propose d'analyser la communication et/ou le langage de différentes populations pathologiques (enfants dysphasiques ou autistes en interaction avec des soignants ou leurs parents). Elle tiendra compte des productions et des comportements de l'enfant mais aussi de ceux de l'adulte.

La recherche procède par comparaison entre documents présentant des similarités au niveau du cadre, de la situation ou des participants. Ces documents sont de plusieurs ordres :

- documents portant sur des situations peu standardisées : enregistrements familiaux, enregistrements d'entretiens thérapeutiques ;
- documents portant sur des situations plus standardisés enregistrées en centre de Protection Maternelle et Infantile dans le cadre d'un test de suivi du regard.

Les différents thèmes retenus sont les suivants :

- i) à partir du corpus de films familiaux recueillis auprès de familles d'enfants autistes, l'analyse portera sur la parole de la mère, les variations du comportement autistique chez l'enfant, les signes précurseurs de la dysphasie, et sur les rythmes des échanges ;
- ii) à partir d'enregistrements vidéoscopés d'entretiens thérapeutiques (enfants autistes, enfants sévèrement dysphasiques), l'analyse portera sur les variations au cours de l'entretien thérapeutique par l'analyse de la mimo-posturo-gestualité de l'intonation et du discours de l'adulte autiste, par l'analyse des échanges de regard entre un enfant autiste et son thérapeute, et par la comparaison de l'accès progressif au langage chez trois enfants dysphasiques profonds. Il s'agit de caractériser la diversité des accès au langage en situation pathologique, de montrer que l'accès au langage chez un enfant dysphasique ne transite pas par les étapes que l'on observe chez l'enfant tout venant.
- iii) à partir d'enregistrements issus d'un protocole de dépistage (en collaboration avec l'équipe du Pr. Moro, Hôpital Avicenne)

Dans différentes PMI, un test standardisé de suivi du regard a été mis en place avec des enfants de 4 mois afin de dépister les enfants en risque d'évitement relationnel ultérieur. Il est fondé sur l'observation comparée du suivi du regard de la mère quand l'enfant est assis face à sa mère dans un baby-relax, ou quand il est assis sans être soutenu. Dans les deux cas, la mère est d'abord en interaction face à l'enfant et lui parle. Puis elle détourne la tête pour regarder

un objet qui se trouve sur le côté de l'enfant. On regarde si l'enfant détourne la tête pour suivre le regard de la mère. Pour les enfants qui ne parviennent pas à suivre le regard de la mère dans l'une des situations, des consultations régulières sont proposées. Certaines sont filmées afin de suivre l'évolution de l'enfant.

## **2) Action communication avant le langage : intonation et posturo-mimo-gestualité dans le développement normal et le développement autistique**

Cette action poursuivra des analyses de corpus audiovisuels longitudinaux pour établir une typologie des gestes et des intonations à visée communicative dans le développement normal, puis en comparaison avec des films familiaux d'enfants autistes. L'analyse des caractéristiques de ces gestes et intonations dans un contexte d'interaction permettra d'établir d'une part les différences entre le développement normal et le développement autistique, tant sur le plan quantitatif que sur le plan qualitatif, et d'autre part, de mettre en relief les éléments favorisant la communication chez l'enfant autiste.

## **3) Action dysphasie : troubles de développement du langage oral**

La dysphasie se caractérise par un trouble de développement du langage oral qu'on ne saurait attribuer à des causes telles que retard mental, anomalies neurologiques, trouble de l'audition, trouble socio-affectif ou comportemental. L'enfant présente d'abord un certain retard, que souvent il ne pourra jamais combler même avec une aide orthophonique.

Les causes de ces troubles sont encore incomprises et des explications parfois contradictoires ont été proposées. On constate que, dans la plupart des cas, la phonologie et la syntaxe sont atteintes de manière plus ou moins simultanée. On trouve aussi des difficultés d'ordres lexical et pragmatique chez beaucoup d'enfants.

L'hypothèse que nous faisons est que les troubles qui caractérisent l'enfant peuvent avoir des origines différentes même dans le cas de manifestations comportementales identiques, dans la mesure où le système linguistique possède ses propres aspects de régularisation et de compensation qui font que tout composant du langage repose en partie sur les autres. Cet état de fait limite la compréhension des troubles et ne permet pas de focaliser sur de bonnes rééducations. Pour éliminer ces problèmes, deux axes de travail sont en cours.

### *Performances cognitives non linguistiques*

Beaucoup de fonctions langagières ont des pendants dans des domaines non linguistiques. Par exemple, on sait que les troubles de mémoire de travail accompagnent souvent mais pas systématiquement les troubles de langage. Il en est de même pour les troubles de traitement rapide des sons. L'évaluation non langagière permettra peut-être de découvrir des profils dans les dysphasies, ce qui permettra d'améliorer et d'affiner les rééducations.

### *Aspects longitudinaux du développement syntaxique*

Les performances des enfants ne sont pas toujours stables au cours du temps. Ainsi les performances phonétiques faibles s'améliorent souvent avec la rééducation sans que le reste du langage progresse pour autant. Un autre exemple est celui des erreurs syntaxiques qui peuvent n'apparaître qu'à un certain moment du développement lorsque le langage de l'enfant est suffisamment développé pour qu'il ne puisse plus se contenter de formules figées. Le suivi longitudinal des enfants dysphasiques en parallèle de celui d'enfants-contrôles devrait permettre de mieux distinguer les différences de troubles. Il permettra aussi de tester en situation des hypothèses comme celle où les constructions complexes de l'enfant ne sont construites qu'à partir de ses productions antérieures, contrairement aux hypothèses mettant en jeu une syntaxe innée.

#### **4) Remédiation à la dyslexie**

L'objectif essentiel est une évaluation du jeu de Play On destiné à favoriser la récupération des enfants dyslexiques, qui existe à présent sous forme de CD-ROM (pour PC). En association avec l'association CORYDIS regroupant des parents d'enfants dyslexiques, une subvention de l'Education Nationale permet de faire appel à des chercheurs afin d'examiner la manière dont les enfants se servent du jeu (enquête de convivialité) puis bénéficient réellement des effets du jeu. L'ensemble de l'expérimentation devrait se faire au cours des deux années universitaires à venir. A l'issue de cette période, outre la publication des résultats des différentes études, nous devrions être à même d'aménager le jeu en fonction des propositions formulées par les équipes d'évaluation.

#### **5) Subjectivité, continuité discursive et troubles d'apprentissage**

L'hypothèse centrale de ce projet réside dans l'interrogation du maniement du langage oral par des enfants présentant des troubles d'apprentissage contrastés (lecture vs mathématiques), afin d'établir si chaque groupe, avec des variations individuelles probables, peut se caractériser par un profil particulier, et lequel. La notion de profil de sujet langagier est à considérer comme une constellation de façons d'être, de raconter et d'interagir langagièrement, qui présente des caractéristiques communes ou analogues ou à l'inverse opposées du point de vue de la subjectivité langagière dans des conduites et des situations diversifiées. De premiers résultats permettent d'émettre une hypothèse sur des caractéristiques langagières différenciées selon que l'enfant présente des troubles d'acquisition de la lecture ou du calcul dans le cadre de l'activité narrative dans un protocole constitué de conduites narratives et argumentatives. Les récits d'enfants en difficulté de lecture se caractérisent par un effacement de la subjectivité langagière, alors que ceux d'enfants en difficulté en mathématiques présentent au contraire un « trop plein » presque envahissant de l'énonciation de la subjectivité. Il s'agit donc d'affiner l'hypothèse en précisant à quels plans du discours et de l'énonciation cet effacement/envahissement s'observent et à quel degré. Cette interrogation sera soutenue par une mise en perspective entre une activité narrative et une activité argumentative programmée et encadrée par un protocole commun et une situation d'entretien de synthèse entre l'enfant, les parents et le médecin qui propose ou conseille une remédiation. La recherche ici proposée s'inscrit dans un contexte social particulier. Comme nous l'a rappelé le rapport INSERM, l'ampleur des troubles sévères des apprentissages est telle que les enfants et les adolescents concernés par une telle situation sont aussi nombreux que l'ensemble de ceux qui souffrent de tous les autres handicaps.

Sous-tendant l'ensemble de la recherche, la théorisation de l'inscription de la subjectivité dans le langage par certains courants de la linguistique contemporaine permet de prendre en compte à la fois les dimensions de l'affectivité et du point de vue. Le langage n'est plus considéré comme distinct de ses utilisateurs qui peuvent imprimer à leur discours la marque de leur point de vue et de leur ressenti. C'est cette inscription dont nous cherchons à évaluer la présence en relation avec la notion de cohérence discursive ainsi qu'avec celle de style énonciatif et de façons d'interagir.

#### **2.b. Résultats attendus**

Sur le plan du dépistage précoce de l'autisme le projet donnera lieu à la publication d'un ouvrage et à une journée d'étude confrontant positions cognitivistes et psychanalytiques.

Sur la plan de la recherche portant sur la dysphasie, on attend une meilleure compréhension des aspects longitudinaux du développement phonologique et morphosyntaxique et en particulier de la différence entre développement pathologique et développement normal sur chacun de ces deux aspects. Par ailleurs, le programme permettra de mieux connaître les liens entre compétence non langagière et développement de la compétence langagière.

En ce qui concerne les travaux sur les troubles d'acquisition, les résultats attendus concernent une différenciation dans le maniement langagier (récits, argumentation du protocole et entretien) des diverses dimensions de la subjectivité langagière, et la mise en évidence d'affinités voire de corrélations entre une spécificité de ce maniement et le type de troubles d'apprentissage.

Une thèse est programmée sur l'émergence de la grammaire chez des enfants présentant des troubles spécifiques du langage oral.



### 3.3.2.2. Equipe *Dynamiques des usages*

**Coordination : Françoise Gadet & Sylvie Plane**

**Membres de l'équipe *Dynamiques des usages***

<b>Membres rattachés à titre principal</b>	Chercheurs et enseignants-chercheurs	M. Baker, M. Carcassonne, D. Cogis, C. Corblin, M. Froment, F. Gadet, N. Lewi-Dumont, C. Pauleau, S. Plane, J.-M. Pottier	10	11
	ITA	F.-X. Bernard	1	
	Membres émérites ou honoraires			
<b>Doctorants</b>	D. Balaci, J.-D. Bellonie, S. Benslimane F. Charles, M. Farhat, F. Favart, M.-F. Fradet, W. Guehria, D. Josèphe, B. Lafourcade, A. Makerova, N. Michelis, I. Mouclier, J. Peuvergne, L. Rais, F. Rondelli, M. Ros-Dupont, J. Sebbar, B. Soh bani		19	
<b>Jeunes chercheurs associés</b>	E. Guerin		1	
<b>Membres associés</b>	J.-P. Jaffré (CNRS), H. Tyne (Southampton)		2	
<b>Collaborations extérieures</b>	D. Alamargot, V. Bourhis, C. Brissaud, P. Cappeau, C. Donahue, I. Fenoglio, Cl. Garcia-Debanc, F. Ganier, J. Harbord, T. Olive, S. Pétilion, W. Oesterreicher		12	

#### **Présentation générale de la problématique scientifique**

L'équipe *Dynamiques des usages* s'est donné ce titre avec l'idée que les termes « dynamiques » et « usages » signalent qu'il s'agit de faire porter les recherches sur des processus évolutifs et constatables, et non de prétendre figer des objets langagiers dans des catégories inertes qui seraient aisément cernables et en nombre fini, comme le sont les termes traditionnellement en vigueur : français populaire, français familier, niveau de langue ; langue des jeunes, français scolaire, oral/écrit... Nous adoptons également cette dénomination car elle permet de recouper des problématiques en vigueur dans des disciplines davantage formalisantes, qui se disent « fondées sur l'usage » (recouplement ne voulant pas dire similitude de point de vue). En outre, l'idée même de « dynamiques », qui est au cœur du questionnement de cette équipe, attire l'attention sur le fait que les recherches menées s'intéressent toutes à des phénomènes inscrits dans une temporalité observable.

Le cadre de travail partagé par les membres de l'équipe prend comme point d'ancrage le fait que le langage est une activité sociale et écologique. De là vient le choix d'une double orientation. Une orientation sociolinguistique, si la référence à cette discipline est prise en un sens très large couvrant à la fois des domaines comme l'étude de la variation et du changement, l'écologie des langues, l'ethnographie de la communication, la politique linguistique, les interactions, l'étude des pratiques et des discours ; et une orientation didactique, qui renvoie aux spécificités langagières des interactions et de la production verbale réalisées à des fins d'apprentissage ou de socialisation scolaire.

Les objets favorisés concernent les pratiques intervenant par le truchement de l'oral comme de l'écrit, prenant place en contexte scolaire ou non, professionnel ou non, ordinaire ou non. Il s'agit tout autant d'observer les pratiques actuelles que l'émergence de nouvelles pratiques.

S'agissant du sujet locuteur ou scripteur, on s'intéresse à dégager à la fois des régularités et des singularités, entre locuteurs, entre groupes, entre communautés.

Le point commun est finalement la prise au sérieux de la notion d'utilisateur (que ce soit comme locuteur ou comme scripteur) en tant qu'acteur, engagé dans des interactions écologiques et sociales situées. Ceci afin de montrer que les objets observés sont toujours d'une extrême complexité qu'on ne cherchera aucunement à réduire mais bien à décrire, à comprendre, voire à expliquer. Nos méthodes sont donc avant tout fondées sur l'empirie : détermination de protocoles de recueil de données, observations, constructions de corpus, modalités de transcription, comparaisons, classements, allant éventuellement jusqu'à la modélisation.

Les travaux de l'équipe « Dynamiques des usages » sont ici organisés en deux opérations, mais elle est aussi impliquée dans les opérations transversales.

Les deux opérations propres à l'équipe s'intéressent à la production langagière, à propos de laquelle elles étudient respectivement l'effet des médiations qui s'exercent sur elles dans le cas des situations d'apprentissage, la variété des styles et des dynamiques de la production écrite et de la production orale.

### **Liste des opérations**

3.3.2.2.1. *Variétés des styles et des dynamiques de la production orale*

3.3.2.2.2. *Variétés des styles et des dynamiques de la production écrite*

#### **3.3.2.2.1. Variétés des styles et des dynamiques de la production orale**

##### **1. Membres de l'opération**

*Membres rattachés à titre principal (8 : 2 PR, 4 MC, 2 DR)*

Françoise Gadet (responsable)

Michael Baker

Marie Carcassonne

Mireille Froment

Christian Hudelot

Marc Klein

Sylvie Plane

Christine Pauleau

*Doctorants (14 dont 6 en cotutelle)*

Diana Balaci

Jean-David Bellonie

Sahar Benslimane

Mokhtar Farhat

Françoise Favart

Wajih Guehria

Danièle Josèphe

Anna Makerova

Nicoletta Michelis

Isabelle Mouclier

Julie Peuvergne

Larbi Rais

Jamila Sebbar  
Baligh Sohmani

*Jeune chercheur associé (1)*  
Emmanuelle Guerin

*Membre associé (2)*  
Jean-Pierre Jaffré (CNRS)  
Henry Tyne (Université de Southampton)

*Collaborations extérieures (2)*  
Paul Cappeau (Université de Poitiers)  
Wulf Oesterreicher (Université de Munich)

## **2. Problématique**

L'opération sur les dynamiques dans la production orale vise à documenter la variété des productions orales (ou du moins, celles dans lesquelles c'est l'oral qui tient une place prépondérante), en les étudiant par le biais des interactions et du dialogisme, plus que des situations elles-mêmes. C'est cette exigence de point de vue sur le dialogue et l'interaction qui dicte les terrains retenus, la plupart d'entre eux étant sélectionnés dans la vie scolaire, mais pas seulement : des pratiques sociales partie prenante de la vie scolaire qui mettent en œuvre la notion de genre discursif, comme par exemple dans l'observation de débats en classe (action 2), ou le fonctionnement des conseils de la vie scolaire (action 3), ainsi que de dialogues scolaires (action 4) ; mais aussi l'amplitude de la variation chez un locuteur unique, que l'on appelle « style » (action 1).

### **2.a. Actions**

#### **1) Documenter la diversité stylistique**

Cette action prend la suite de l'action sur le style (voir partie Bilan), cherchant à documenter la diversité des productions orales en les abordant à travers cet angle d'attaque privilégié, qui tire les conséquences de ce que la diversification stylistique est une propriété des langues dans leur usage, et que le style est un phénomène définitoire des langues, permettant un accès à leurs dynamiques. Le point de départ est dans le locuteur, dont les « façons de parler » ne sont pas passivement soumises aux situations dont il est partie prenante, mais qui dispose d'une certaine latitude d'initiative, qu'il met en œuvre dans la palette diversifiée de ses ressources linguistiques et langagières.

Les objets pris en compte peuvent être très divers, puisque l'objectif est de documenter tout événement de langage qui pourra venir nourrir la réflexion sur la variation stylistique et sa place dans la représentation des langues et leurs frontières. Pour le moment, différentes thèses en cours se fixent sur les objets suivants : la façon dont l'oral met en jeu un rapport à l'écrit (thèses Bellonie, Favart, Josèphe), la maîtrise du style en L2 (thèses Michelis, Guehria), la langue des jeunes (thèses Balaci, Bellonie, Makerova, Sohmani), le style dans différentes situations de francophonie (thèses Benslimane, Guehria, Peuvergne, Rais, Sebbar) et professionnelles (thèse Mouclier), l'exploitation de l'oralité chez les comiques (thèse Farhat).

## **2) Le débat à l'école : autour du thème du rapport à l'autre, construction des points de vue et hétérogénéité des sources énonciatives**

Le projet porte sur *la construction des points de vue et l'hétérogénéité des sources énonciatives* dans des débats en classe ayant pour thème le rapport à l'autre. En effet, les récentes Instructions Officielles pour l'école primaire incitent à développer les capacités argumentatives des élèves dans des débats au cours desquels il est recommandé de « faire appel à leur expérience personnelle ».

La pratique du débat scolaire travaille dans une « zone de contact », entre le monde de l'école et celui du hors école, où se trouvent confrontées différentes normes et valeurs, diverses façons de dire et de faire. Les pratiques concrètes dans la classe constituent les observables de cette hétérogénéité, de même que des tensions et contradictions qu'ils génèrent.

L'école, tout en étant un lieu de préconstruit social et culturel, est traversée par un écheveau complexe de points de vue hétérogènes et plus ou moins compatibles portant aussi bien sur les savoirs et l'apprentissage que sur les normes et les valeurs. Les acteurs peuvent être pris dans des contradictions, des tensions plus ou moins aiguës du fait de la diversité des façons d'être, de dire et de faire attachée à l'hétérogénéité des groupes de socialisation, ce qui ne peut pas ne pas avoir d'effets sur la construction des points de vue.

Les données sont constituées de l'enregistrement audio de débats recueillis dans des classes de CE1/CE2 (8-9 ans) d'écoles primaires. L'étude vise à éclairer les effets du cadre scolaire et disciplinaire (instruction civique, atelier philo) sur « l'objet dont on parle », sur la façon dont se construisent les points de vue sur un thème. On écarte la conception d'un sujet locuteur conscient et unifié. La question du point de vue, de sa construction et de ses évolutions conduit à prendre en compte la diversité des voix, la relation entre « construction des points de vue » et « sources énonciatives », c'est-à-dire hétérogénéité des énoncés, des capacités énonciatives, des sources de connaissances.

C'est l'enchaînement textuel qui montre l'intention discursive, non la structure de l'énoncé. Bakhtine propose une conception large du contexte culturel, créée par le discours d'autrui : en termes d'opinions, de jugements de valeur, de points de vue. L'objet de l'étude est alors le dialogisme qui structure le discours (mode discursif, reprise-modification du discours d'autrui, conduites de métaphore et de mélange).

Pour penser concrètement ces problèmes, il faut élaborer une méthodologie spécifique qui permette d'observer comment se construisent les points de vue, dans le milieu scolaire et dans le cadre de l'hétérogénéité des sources énonciatives. Plutôt que de privilégier la présence ou l'absence de marques de la diversité des plans énonciatifs et des modes de prise en charge des énoncés par un locuteur, nous nous intéressons aux mouvements dans le déroulement des discussions, à la dynamique temporelle des dialogues, à l'entrelacs entre genre, thème, place et manifestations idéologiques, non assignables dans tel mot ou telle construction syntaxique.

## **3) Événements discursifs et travail langagier dans les conseils d'élèves**

Elaboration d'un modèle d'analyse destiné à rendre compte des événements discursifs et du travail cognitivo-langagier qui s'opère à l'occasion de la tenue des conseils de « vie scolaire ». Ces conseils réunissant élèves et enseignants constituent l'une des réponses institutionnelles aux phénomènes d'incivilité et de violence qui affectent le milieu scolaire. Ces conseils ont pour objectif de donner aux élèves et aux maîtres l'occasion de revenir sur des comportements, des ressentis et des points de vue plus ou moins spontanés, pour entrer dans une démarche réflexive et pacifier les relations interpersonnelles. Les Conseils ouvrent la possibilité de *débattre* et de *légiférer* à propos des rapports entre pairs, entre individu et institution, de la vie sociale au sens large du terme. De ces pratiques concrètes dépendent la

qualité des relations interpersonnelles et plus spécifiquement la façon dont se joue le rapport de rôle enseignant-élèves.

Dans les conceptions qui ont présidé à la mise en place de ces conseils, la verbalisation et le dialogue sont considérés comme un moyen d'analyser, d'équilibrer les tensions, d'éviter le recours à l'acte violent, en favorisant l'explicitation des implicites scolaires et relationnels.

On confrontera l'efficacité pour interpréter les faits langagiers collectés de deux modes d'analyse, relevant pour l'un de la linguistique interactionnelle et de l'analyse conversationnelle, et pour l'autre d'une approche interprétative qui tient compte de l'hétérogénéité constitutive du sujet parlant.

#### **4) Dialogue, accordage et temporalité dans le dialogue scolaire**

Dans le cadre global de l'organisation de la parole en interaction dans le contexte scolaire, on se propose d'analyser la spécificité des interactions maître-élèves et/ou élèves-élèves, du point de vue de la temporalité du discours d'un sujet dans sa relation à la temporalité du discours d'un autre sujet (d'autres sujets), et en tenant compte des aspects non verbaux (posturo-mimo-gestuels). Il s'agit d'aborder cet aspect de l'interaction par le biais du lien subjectivité et temporalité. La temporalité du discours d'un sujet dans sa relation avec celle du discours d'un autre : temporalité du discours de la maîtresse en relation avec la temporalité du discours des élèves, temporalité du discours d'un élève en relation avec celle d'un autre élève et/ou des autres élèves.

#### **2.b. Résultats attendus**

Différentes communications (colloques ou conférences invitées) sont programmées, conférences pour les doctorants, organisation de journées d'étude, réalisation d'un numéro de revue (revue généraliste, et revue didactique), rédaction d'articles sur des situations spécifiques, monographies. Projet d'ouvrage : Mireille Froment (*Les Conseils : Substituer la rationalité discursive au geste violent ?* titre provisoire).

#### **3.3.2.2.2. Variétés des styles et des dynamiques de la production écrite**

##### **1. Membres de l'opération**

*Membres rattachés à titre principal (6 : 2 PR, 4 MC)*

Sylvie Plane (responsable)

Danièle Cogis

Colette Corblin

Françoise Gadet

Nathalie Lewi-Dumont

Jean-Michel Pottier

*Doctorants (5 dont 1 en cotutelle)*

Diana Balaci

Jean-David Bellonie

Marie-Françoise Fradet

Bernadette Lafourcade

Michelle Ros-Dupont

*Jeune chercheur associé (1)*

Emmanuelle Guerin

*Collaborations extérieures (9)*

Denis Alamargot (LMDC, Université de Poitiers)

Catherine Brissaud (LIDILEM, IUFM de Grenoble)

Véronique Bourhis (IUFM de Bretagne)

Irène Fenoglio (ITEM, CNRS)

Claudine Garcia-Debanc (IUFM de Midi-Pyrénées)

Franck Ganier (Université de Haute-Bretagne)

John Harbord (Université de Budapest)

Thierry Olive (LMDC, Université de Poitiers)

Sabine Pétilion (ITEM, CNRS)

## **2. Problématique**

L'objectif de cette recherche est de fournir des éléments pour déterminer des profils de scripteurs afin d'affiner les modélisations du processus scriptural dont on dispose actuellement, qui sont dérivées pour la plupart du modèle princeps de Hayes & Flower 1980. Ces modélisations formelles n'intègrent ni les caractéristiques idiosyncrasiques de l'activité scripturale, ni les spécificités génériques des textes à produire. En outre, elles s'intéressent peu à la manière dont s'opère la textualisation proprement dite. Il s'agit donc de collecter des données exemplaires et de les analyser afin de faire apparaître l'étendue des variations qui entrent en œuvre dans le processus scriptural, et de décrire les manières dont s'opère la genèse et le tissage textuel. Cette recherche s'intéresse plus particulièrement à deux familles de genres d'écrits, les écrits de synthèses et les écrits narratifs dont la production recouvre des enjeux scolaires.

Du point de vue méthodologique, les données sont recueillies soit en situation écologique soit en situation expérimentale. Elles sont de trois ordres : traces graphiques, enregistrements d'interactions verbales lors de l'écriture collaborative, et enregistrements des mouvements oculaires (dispositif *eye and pen* de Alamargot & Chesnet). Dans tous les cas, la recherche a recours à un même cadre théorique permettant de traiter de façon cohérente les données recueillies à partir des outils d'analyse issus des trois champs de recherche s'intéressant à la textualisation : linguistique textuelle, psycholinguistique de l'écrit, et critique génétique.

Les variables examinées pour déterminer des profils de scripteurs concernent : a) les modes de traitement scriptural de la consigne, b) les modes de gestion des contraintes linguistiques, c) les modes d'appropriation de l'espace graphique en fonction de l'outil graphique, et d) les caractéristiques temporelles du processus scriptural.

### **2.a. Actions**

#### **1) Traitement des contraintes et variétés des modes rédactionnels**

Cette recherche s'attache à repérer des variables et des invariants dans les modes rédactionnels, de façon à discerner comment s'opère le développement des capacités rédactionnelles.

Pour des raisons d'ordre méthodologique et théorique, on considère la production d'écrit comme une activité au cours de laquelle le scripteur est amené à traiter à la fois des contraintes d'ordre psycholinguistique, dues aux limites de ses capacités mémorielles, et des contraintes d'ordre linguistique de niveaux inégaux mais non hiérarchisables. La complexité de ces facteurs fait que les processus scripturaux connaissent nécessairement de grandes variations dans leur mise en œuvre, mais aussi des spécificités dues aux caractéristiques partagées par des scripteurs relevant de catégories relativement homogènes en termes d'âge ou d'habiletés linguistiques.

La recherche comporte deux volets, Le premier volet s'attache à décrire et comparer des modes de textualisation en s'attachant à la manière dont s'opère le traitement de contraintes éventuellement antagonistes, ainsi qu'à l'inscription de ce traitement dans la temporalité de l'écriture.

Le deuxième volet vise à rendre compte des différentes formes d'évolution de l'activité rédactionnelle, envisagée non pas uniquement sous l'angle des performances, mais aussi sous celui de l'activité du scripteur, de façon à mettre en évidence les faisceaux de causalités qui s'exercent sur cette évolution.

Les deux volets de cette recherche s'appuient sur l'analyse comparée de protocoles d'écriture auxquels sont soumis des scripteurs différenciés en fonction de critères spécifiques.

## **2) Aide à l'écriture et traitement de la consigne**

Le projet de recherche porte sur la comparaison d'écrits répondant à une consigne identique avec un script imposé, mais utilisant deux aides d'écriture différents. La méthodologie est celle de l'analyse comparative de productions d'écrits répondant à des situations différentes.

L'étude porte sur la production d'un récit en classe de CM2 ayant pour thème la narration d'un voyage scolaire fictif ; d'une part, on fournit au premier groupe de scripteurs un dessin d'itinéraire légendé par des indications sommaires concernant les horaires et les lieux, d'autre part on fournit au deuxième groupe une liste écrite renvoyant à ces horaires et à ces lieux.

L'objectif est de déterminer d'abord si le type d'aide à la narration (schéma *vs* texte) contraint le style de la narration, l'emploi des temps et des connecteurs spatiotemporels en particulier. Il s'agit donc de définir par quoi se manifeste l'éventuelle contrainte exercée sur l'écriture. La visée est pédagogique, puisqu'on cherche à définir comment l'aide à l'écriture, par sa nature, langagière ou non, pourrait avoir un impact sur le développement des compétences narratives, et pour quels enfants.

## **3) Modes de collaboration dans l'écriture dite « à quatre mains »**

Cette recherche porte sur les modes de collaboration dans l'écriture collaborative. Elle s'intéresse plus particulièrement à l'écriture dite « à quatre mains », qui recouvre des modalités de rédaction fréquentes et diverses dans le domaine du littéraire à partir de 1850.

Le volet épistémologique de cette recherche concerne l'évolution des modalités d'écriture. Il s'agit de définir le cadre méthodologique d'un recueil de données significatif permettant de traiter l'hétérogénéité des éléments d'information issus des traces graphiques et des mises en scène littéraires de l'activité d'écriture, et de constituer un corpus rassemblant des exemples littéraires d'écriture collaborative, des entretiens et des témoignages afin de faire un historique retraçant les évolutions historiques de l'écriture collaborative et de sa réception.

Le deuxième volet de cette recherche porte sur l'analyse de protocoles d'écriture collaborative différenciés selon des variables déterminées par l'analyse épistémologique.

#### **4) Dynamique d'acquisition de l'orthographe**

Dans le prolongement du quadriennal 2004-2007, qui avait donné lieu à une enquête importante sur les compétences orthographiques des élèves, cette recherche a pour but d'approfondir l'identification des représentations des élèves dans le domaine de la morphographie, afin d'affiner le modèle actuel en menant des entretiens avec des élèves de collège et de valider les propositions didactiques. Cette recherche en cours fait l'objet d'une collaboration étroite avec Catherine Brissaud, du laboratoire LIDILEM de l'Université de Grenoble 3.

#### **5) Gestion des contraintes orthographiques en production écrite**

Cette recherche pluridisciplinaire s'inscrit dans le cadre du GDR « Pratiques de production Verbale Écrite » (2007-2011) (Responsable : Denis Alamargot, MC IUFM Poitou-Charentes, Laboratoire LMDC-CNRS : Gestion des contraintes cognitives et linguistiques).

La premier volet de la recherche vise à évaluer l'impact des contraintes orthographiques sur le processus de production écrite et l'effet sur l'acquisition et le développement, en croisant les approches psycholinguistiques et linguistiques. Pour décrire la variation rédactionnelle en fonction du contexte d'écriture, diverses variables sont examinées : expert-novice, genres discursifs contrastés (narratif/explicatif), production initiale-révision, caractéristiques orthographiques différenciées finnois/français (orthographe transparente vs opaque).

Un second volet de cette recherche s'intéresse plus particulièrement à la description des dynamiques de scription. L'analyse en temps réel des interventions de différents scripteurs permet de déterminer le mode de prise en compte des normes orthographiques en cours d'écriture.

#### **2.b. Résultats attendus**

Cette opération centrée sur la description de l'activité scripturale s'inscrit dans la perspective d'un dialogue interdisciplinaire entre spécialistes de sciences du langage et spécialistes de psychologie cognitive. Des publications conjointes sont prévues un numéro de la revue *Langages* portant sur le traitement des contraintes linguistiques et psycholinguistiques de la production d'écrit et un ouvrage collectif présentant l'activité de textualisation d'un triple point de vue, linguistique, psycholinguistique et didactique.



### 3.3.3. OPERATIONS INTER-EQUIPES

Nous avons souligné dans notre déclaration de politique scientifique l'importance que nous accordions aux synergies internes au laboratoire. C'est dans ce cadre que se sont structurées quatre opérations inter-équipes.

D'une part, deux opérations impliquant chacune une équipe de la composante Modèles et une équipe de la composante Dynamiques : dans la mise en relation entre la langue comme appropriation des locuteurs et l'objet qu'ils se construisent et peuvent dès lors manipuler, comme d'autres formes de savoirs, il était important de mettre en relation, dans des opérations trans-composantes, les dynamiques langagières et les modèles linguistiques et/ou discursifs.

D'autre part, deux opérations, structurées entre les deux équipes de la composante Dynamiques des langues, dans ce qu'elles ont de complémentaire : les dynamiques des acquisitions et les dynamiques des usages, plus proches de la psychologie pour la première équipe, de la sociologie ou de l'ethnologie pour la seconde. La porosité de la frontière se lit à la circulation de chercheurs entre les deux équipes.

#### **Liste des opérations**

3.3.3.1. *Corpus de langage oral : formats, conversion, visualisation et interrogation*

3.3.3.2. *Variétés des genres discursifs scolaires et universitaires*

3.3.3.3. *Dynamiques acquisitionnelles et sociolinguistiques des situations de la francophonie périphérique*

3.3.3.4. *Interactions langagières et acquisitions*

#### **3.3.3.1. Corpus de langage oral : formats, conversion, visualisation et interrogation**

##### **1. Membres de l'opération**

*Membres rattachés à titre principal (6 : 3 PR, 1 DR, 1 CR, 1 IR)*

Christophe Parisse (responsable)

Françoise Gadet

Christian Hudelot

Anne Lacheret

Jean-Luc Minel

Colette Noyau

##### **2. Problématique et description**

Pour rentabiliser les efforts consentis pour constituer des corpus, il faudrait disposer d'outils qui permettent de les exploiter au mieux, y compris avec des approches pour lesquelles ils n'étaient pas nécessairement conçus à l'origine. Le but principal de l'opération est de créer de tels outils utilisables sur des données de toute langue et qui complètent l'offre logicielle déjà disponible.

## **2.a. Actions**

Les outils que nous entendons développer permettront de valoriser les résultats des autres actions du laboratoire. Les outils multi-plateformes (Windows, Macintosh, Linux) développés utiliseront uniquement des logiciels gratuits et « open-source » et seront rendus disponibles à la communauté scientifique sous la même forme, ce qui garantit la possibilité d'extension future. Cette ouverture est également assurée par l'utilisation de codages normalisés (XML et Unicode) sous contrôle de l'expertise du CRDO (Michel Jacobson) et compatibles avec ceux qui sont développés actuellement dans les projets CHILDES et PHON.

Les outils disposeront d'interfaces graphiques conviviales, et permettront l'accès à distance à des corpus mutualisés. Des procédures de conversion de format seront réalisées pour les logiciels les plus utilisés pour la transcription de corpus de langage oral. Trois types d'outils seront développés :

### *a) Conversion*

Les corpus de langage oral sont souvent issus de logiciels différents et leur format varie énormément. Le but de l'action est de fournir des outils permettant de convertir et normaliser les corpus existants vers un format commun ouvert et utilisable par tous les logiciels développés dans le laboratoire.

### *b) Interrogation*

L'interrogation de corpus annotés doit pouvoir porter sur tout matériel figurant dans les transcriptions. L'interrogation sera du type base de données et pourra porter sur une partie de document (par exemple, un locuteur spécifique, les interactions qui précèdent ou suivent une situation donnée, etc.), sur un document complet ou sur un groupe de documents (par exemple, tous les enregistrements concernant un locuteur cible, ou correspondant à un type d'enfant précis, à un âge, à un sexe, à une situation de recueil, etc.). Les résultats d'interrogation seront dans les mêmes formats que les corpus afin de pouvoir être facilement visualisés ou édités.

### *c) Visualisation des corpus et des résultats d'interrogation*

Comme les corpus peuvent comporter des informations de niveaux très différents, des locuteurs ou situations variées, ainsi que des annotations sonores ou vidéo, l'ensemble d'une transcription est souvent difficile à lire. C'est pourquoi il est nécessaire de disposer d'outils de visualisation permettant de ne présenter que les éléments souhaités avec des formats graphiques appropriés (présentation en ligne ou en colonne, utilisation de marques typographiques ou de couleurs, images, sons, vidéo, repères temporels, etc.).

## **2.b. Résultats attendus**

Ce projet transversal à l'ensemble du laboratoire vise à inciter tous les membres à mieux formuler leurs besoins en matière de traitements de corpus oraux, sur la base d'un véritable va-et-vient : exprimer des demandes précises, tester les outils proposés, faire un renvoi de satisfaction et éventuellement des demandes de précisions. On envisage de mettre l'expérience ainsi accumulée à la disposition de membres d'autres laboratoires, en France ou à l'étranger.

### **3.3.2.2. Variétés des genres discursifs scolaires et universitaires**

#### **1. Membres de l'opération**

*Membres rattachés à titre principal (4 : 3 PR, 1 MC)*

Sylvie Plane (responsable)

Catherine Boré

Colette Corblin

Marie-Laure Elalouf

*Doctorants (2)*

Florence Charles

Fabienne Rondelli

*Collaboration extérieure (1)*

Christiane Donahue (University of Farmington, Maine)

#### **2. Problématique**

Ce projet de recherche rassemble des chercheurs engagés dans l'équipe *Activité langagière* de la composante Modèles, et dans l'équipe *Dynamiques des Usages* de la composante Dynamiques, qui, les uns et l'autres s'intéressent à la production d'écrits telle qu'elle est effectuée dans le cadre scolaire ou universitaire. Il est construit en lien avec le réseau européen COST European Research Network on Learning to Write Effectively (ERN-LWE), dont il constitue un axe.

Le projet de recherche porte sur la caractérisation de genres d'écrits en usage dans l'enseignement scolaire et universitaire et sur la mise en évidence des variations qui les affectent en fonction des contextes d'enseignement dans lesquels ils sont pratiqués.

Les genres d'écrits scolaires et universitaires sont organisés en un petit nombre de grandes catégories regroupées sous des dénominations génériques. Ils sont définis par des règles explicites qui déterminent les critères officiels des évaluations institutionnelles auxquelles ils sont soumis. Cependant, en marge de cette caractérisation formelle qui ne rend compte que très imparfaitement de la réalité des fonctionnements des écrits effectivement produits, il existe des règles implicites propres à des communautés discursives données. Ces règles se sont constituées empiriquement à partir de traits fonctionnels fournis par l'accumulation des exempla et sont donc incomplètement figées, mais pleinement opératoires. En particulier, on constate que sous les vocables communs de « résumé », « synthèse », « texte argumentatif » et « texte explicatif », les caractéristiques attendues des textes produits diffèrent selon les niveaux, les secteurs et les lieux d'enseignement.

L'objectif de cette recherche est donc de mettre au jour les règles implicites d'écrits dont la production est exigée dans des évaluations scolaires et universitaires et d'identifier les variables non codifiées qui régissent de fait les jugements portés sur ces écrits et l'évolution de ces variables.

Trois méthodologies complémentaires seront employées :

- Analyse comparative des prescriptions relatives à ces exercices ;
- Analyse comparative de productions d'écrits réalisées dans des contextes différenciés ;
- Tâches de jugement.

Les résultats de cette recherche devraient contribuer à l'amélioration des mesures européennes d'harmonisation des dispositifs d'évaluation des enseignements, et en particulier permettre d'assurer une meilleure fiabilité des grandes enquêtes du type PISA.

## **2.a. Actions**

### **1) Linguistique des textes et écriture scolaires**

L'objectif à long terme de ce projet est de caractériser linguistiquement l'écriture scolaire et les genres de discours qui la constituent, en partant de l'hypothèse que les genres sont les outils par lesquels s'enseigne l'écriture.

La production écrite scolaire constitue en effet un objet didactique mal (ou peu) caractérisé. De nombreux chercheurs en didactique et en linguistique, s'appuyant sur les travaux de Bakhtine, mettent en évidence le fait qu'il s'agit de genres d'écrits intermédiaires entre les « genres premiers » et les « genres seconds ». Ces travaux se fondent également sur les thèses de Vygotsky sur l'interactionnisme social, pour proposer un modèle de la production écrite construit sur les interactions langagières.

Notre recherche a permis de circonscrire deux types de problèmes :

a) Extension et définition de la notion de « contexte » liée à la production de types de textes scolaires. Il semble nécessaire de fournir une analyse fine des interactions orales inter-élèves et élèves-enseignant dans la production des genres.

b) Analyse des « genres scolaires » produits

Nous disposons actuellement de trois corpus génériques homogènes qui ont été réunis en classe de 6ème et en fin d'école primaire, de « l'état naissant » à la copie finale.

Ces corpus répondent à une conception du texte ouverte et évolutive, selon laquelle l'écriture est envisagée comme processus observable et comme pratique sociale contextualisée. L'utilisation du logiciel MEDITE (Fenoglio, Ganascia, Lebrave) d'analyse automatique des brouillons est en cours d'expérimentation. Ce travail porte sur des unités linguistiques et des zones de textes limitées dans les productions scolaires, mais jugées significatives dans la définition des genres : incipits, excipits, temps verbaux, dialogues.

### **2) Prescriptions officielles et standards implicites des écrits à visée d'apprentissage ou d'évaluation**

L'objet de la recherche est de caractériser des genres d'écrits en usage en milieu scolaire ou universitaire, utilisés soit à des fins d'apprentissage soit comme instruments d'évaluation. Il s'agit de faire apparaître la diversité des pratiques et des attentes concernant un même écrit, en fonction du contexte scolaire, social ou culturel et de la classe d'âge des élèves.

Cette recherche a pour finalité de fournir des éléments favorisant la cohérence des évaluations, en particulier des évaluations institutionnelles, et de clarifier les standards afin de permettre des comparaisons internationales notamment dans le cadre d'enquêtes du type PISA.

Ce projet prend comme point de départ le fait que certains types d'exercices écrits, tels le *résumé* et l'*explication*, sont pratiqués à tous les niveaux de la scolarité, depuis l'école primaire jusqu'aux examens de niveaux élevés, dans l'ensemble des pays occidentaux. Outre leur usage scolaire, ces deux exercices constituent en général également une pratique scripturale en usage dans le cadre de la communication ordinaire ou professionnelle. Cependant la fréquence de leurs emplois, la multiplicité des fonctions qu'on leur affecte, la diversité des contextes linguistiques et culturels dans lesquels on les pratique, font que les termes génériques comme celui de « résumé » recouvrent des productions scripturales relativement diversifiées.

Il importe donc de mettre au jour les variables qui participent de la définition de ces objets langagiers tels qu'ils sont effectivement mis en œuvre dans différentes institutions ou à différents niveaux de scolarité, et de rendre visibles les attentes implicites qui leur sont liées et

qui sont parfois en concurrence avec les attentes explicites définies par les prescriptions institutionnelles. La méthodologie employée recourt à l'analyse comparative des prescriptions et à des tâches de jugement portant sur des productions réalisées dans des contextes différenciés.

### **3) Evaluation institutionnelle des écrits narratifs et jugement de réussite**

Le projet de recherche s'inscrit dans une contribution à la recherche sur la caractérisation des genres attendus à l'école et au collège. Il s'agit en particulier de l'analyse d'écrits narratifs de fiction contraints par une amorce. La recherche utilise un support développé par le Ministère de l'Education Nationale dans le cadre des actions d'évaluation proposées aux enseignants. L'élève doit écrire un récit à partir d'une amorce, et d'un script tenant lieu d'aide à l'écriture. L'objectif de cette recherche est de confronter les items relatifs à l'évaluation proposée dans le document ministériel aux jugements de réussite attribués par des enseignants. Peut-on dégager une base de compétences narratives scolaires au cycle 2, au cycle 3 et au cycle d'orientation ? L'ensemble s'inscrit dans une réflexion sur les évaluations scolaires et leur éventuelle harmonisation en Europe. La visée est pédagogique, puisque l'on cherche à observer de manière critique les items des évaluations de l'institution, à différents moments du développement des compétences narratives.

La méthodologie est celle de l'analyse comparative de productions d'écrits réalisés à des niveaux scolaires différents.

### **3. Résultats attendus**

Les travaux menés au sein de cette opération fourniront les instruments permettant de construire des comparaisons et des évaluations internationales, contribuant ainsi à la politique de concertation européenne dans le domaine de l'éducation. Pour confronter et mettre en cohérence les résultats d'équipes européennes travaillant sur des thématiques convergentes une journée d'études européenne est prévue en partenariat avec des membres du réseau européen COST European Research Network on Learning to Write Effectively (ERN-LWE), dont les actes seront publiés dans un ouvrage collectif.

Par ailleurs, est programmé dès cet automne le début de deux nouvelles thèses sur les variétés des dynamiques de la production écrite et sur les stratégies cognitives mises en œuvre dans les productions écrites.

#### ***3.3.3.3. Dynamiques acquisitionnelles et sociolinguistiques des situations de la francophonie périphérique***

##### **1. Membres de l'opération**

*Membres rattachés à titre principal (5 : 4 PR, 1 IR)*

Françoise Gadet (responsable)

Colette Noyau (responsable)

Danièle Flament

François Muller

Atanas Tchobanov

*Doctorants (5 dont 2 en cotutelle)*

Jean-David Bellonie (français de Martinique)

Olivia Binganga (français du Gabon)

Guilhem Florigny (français de l'Île Maurice)  
Wajih Guehria (français d'Algérie)  
Julie Peuvergne (corpus de français d'Afrique)

*Membre associé (1)*

Leila Caïd (Ministère de l'Éducation nationale, La Réunion)

## **2. Problématique**

L'opération vise à modéliser la diversité des situations du français à travers le monde, en prenant comme entrée privilégiée l'acquisition de cette langue, compte tenu des contextes spécifiques de la francophonie qui, partout dans le monde sauf en France, est minoritaire au niveau de l'État. Il s'agit ainsi de relier ces modélisations aux tendances linguistiques se manifestant dans chaque contexte, en vue d'expliquer les dynamiques actuelles affectant la langue française dans chacun d'eux. Elle s'inscrit dans la suite du projet AUF-Campus-Cognitique 2001-2005 qui avait mis en place cette réflexion associant acquisition et sociolinguistique pour un terrain délimité de l'Afrique subsaharienne (cf. section 1.2.2.1 du bilan), dont elle constitue un élargissement.

La modélisation envisagée cherchera à établir un modèle se plaçant au carrefour de facteurs géolinguistiques, sociolinguistiques et acquisitionnels : facteurs de l'acquisition tels que présentés par Klein 1989 (accès à la langue, propension à apprendre, capacité linguistique) en tenant compte des dispositifs d'apprentissage institutionnel en vigueur (ou d'ailleurs de l'absence ou de la faiblesse de tels dispositifs) ; conditions géolinguistiques et sociolinguistiques, telles que représentées par exemple par le modèle du français zéro de Chaudenson (Chaudenson *et al.* 1993). La modélisation spécifiera également les modalités de plurilinguisme résultant des conditions d'usage et d'acquisition.

Les domaines linguistiques à explorer ont pour objet central la syntaxe et le discours, dans leur lien au lexique. On commencera par les zones privilégiées pour la variation que sont la zone verbale (temps-aspects-modes, auxiliaire, infinitif, périphrases verbales, constructions et valences verbales), les pronoms (clitiques et pronoms forts) ou la préposition (préposition à régime et préposition orpheline) ; on traitera également des séquences propositionnelles complexes. On s'intéressera aussi à des phénomènes plus spécifiques d'une ou plusieurs situations données, surtout en contexte de langue seconde.

### **2.a. Actions**

#### **Corpus**

- Français des Îles anglo-normandes : Mari Jones (Université de Cambridge).
- Français acadien : Gisèle Chevalier (directrice du CRLA), Annette Boudreau, Louise Beaulieu, Professeurs à l'Université de Moncton (Nouveau Brunswick) ; corpus du CRLA. G. Chevalier et A. Boudreau sont toutes deux titulaires d'une subvention du CRSH canadien.
- Français ontarien : Raymond et Françoise Mougeon, Professeurs à l'Université York, Campus de Glendon, Toronto ; corpus du Centre de Recherche sur les Contacts Linguistiques (CRCL) dont une copie (son et transcriptions) est confiée à *MoDyCo* par convention.
- Français de Louisiane : Ingrid Neumann-Holzschuh, Professeur à l'Université de Regensburg (Allemagne).

- Français mauricien : Rada Tirvassen, Professeur au Mauritius Institute of Education
- Français de la Réunion : Gudrun Ledegen, maître de conférences à l'Université de la Réunion, Leila Caïd, chercheur associé *MoDyCo*, habilitation en cours.
- Français du Cameroun : Louis-Martin Onguene-Essono, Professeur à l'Université de Yaoundé.
- Français du Gabon : Franck Idiata, Professeur à l'Université Omar Bongo, Libreville.
- Français du Sénégal : Moussa Daff, Professeur à l'Université Cheikh Anta Diop, Dakar.
- Français du Togo ; corpus projet CORUS.
- Français de Côte d'Ivoire : Béatrice Akissi Boutin, chercheur au CLLE (Cognition, Langues, Langage, Ergonomie) / ERSS, Université Toulouse Le Mirail) et ILA (Institut de Linguistique Appliquée), Université Cocody-Abidjan, Côte d'Ivoire.
- Français de Tunisie : Chiraz Anane, enseignant-chercheur, Institut Supérieur des Sciences Humaines, Univ. de Tunis-El Manar.
- Mutualisation de corpus existants : Stefan Pfänder, Professeur à l'Université de Fribourg, responsable du projet CIEL\_F.

La structure de cette opération demande une coordination des actions à plusieurs niveaux et en plusieurs temps :

- proposition et validation par circuit de discussion électronique, à partir de nos réflexions méthodologiques sur les corpus oraux, des rassemblements de données pertinentes pour les comparaisons envisagées, soit par extraction et mise en forme de corpus disponibles, soit par collecte de données spécifiques de façon coordonnée ;
- recueil de données sur certains des terrains, et établissement de normes d'exploitation des corpus préalables et nouveaux, sur le plan théorique et technologique ;
- travaux *MoDyCo* – partenaires individuels pour élaborer des monographies permettant de valider la démarche scientifique à généraliser ;
- mise en convergence des acquis des différents partenaires sur les domaines linguistiques prioritairement retenus, par échanges documentaires et journées d'étude centrées sur un domaine ou une aire géolinguistique déterminée ;
- mise en route d'un ouvrage global, via des journées d'étude comparatives ;
- et enfin, élaboration d'un volume d'ensemble.

NB : Les conditions de l'enseignement du français faisant partie des facteurs envisagés, nous aurons à aborder de façon critique aux plans acquisitionnel et sociolinguistique, sur demande des partenaires, les pratiques didactiques en usage sur différents terrains parmi ceux qui sont envisagés.

## **2.b. Résultats attendus**

Cette opération vise avant tout à éclairer les phénomènes évolutifs affectant les français périphériques par l'examen des conditions de reproduction de la francophonie minoritaire à travers l'école et les facteurs acquisitionnels, compte tenu de l'extrême diversité des conditions sociolinguistiques de son usage. Elle escompte ainsi élucider les facteurs de transmission et d'usage induisant les dynamiques en cours des français périphériques. Un dernier objectif est plus directement grammatical : il s'agit de documenter, en syntaxe, l'extension des potentialités de ce qui continue à être regardé comme la langue française, et éventuellement d'en extraire des généralisations.

La matérialisation des travaux se déclinera sous diverses formes : organisation de plusieurs journées d'étude, conférences, articles sur une situation spécifique, monographies éventuelles, colloque international et volume final.

Des retombées didactiques et éducatives sont à attendre également.

### **3.3.3.4. Interactions langagières**

#### **1. Membres de l'opération**

*Membres rattachés à titre principal (6 : 3 MC, 2 DR, 1 IR)*

Michael Baker (responsable)

François-Xavier Bernard

Marie Carcassonne

Mireille Froment

Christian Hudelot

Catherine Le Cunff

*Membre associé (1)*

Rodica Ailincă (IUFM de Guyane)

*Collaborations extérieures (7)*

Jerry Andriessen (Utrecht University)

Isabelle Dumez-Féroc (IUFM de Poitiers)

Jean-Jacques Girardot (Ecole Nationale Supérieure des Mines)

Maria Beatrice Ligorio (Université de Salerne)

Vittorio Scarano (Université de Salerne)

Claire O'Malley (Université de Nottingham)

Michiel Klønhammer (Directeur ICATT Interactive Media Amsterdam)

#### **2. Problématique**

Cette opération vise à donner une visibilité au fait que l'activité langagière est à la fois activité dirigée vers un interlocuteur (qui peut être éventuellement le sujet lui-même) à propos d'un objet du monde, produite dans le cadre d'une activité commune. Les différentes actions, par delà les diversités inévitables des points de vues adoptés dans un domaine où les sciences du langage ont encore à consolider des paradigmes communément admis, présentent néanmoins un ensemble de caractéristiques communes.

Toutes prennent en compte la complexité des interactions tant sur le plan des productions verbales que sur celui de l'ensemble des éléments qui contribuent à la compréhension/interprétation des messages, que ce soit au niveau des expressions corporelles qu'à celui de l'ensemble des autres artefacts présents physiquement dans le cadre de la communication.

Tous prennent également en compte le fait que les interactions langagières ne sont pas sans effet sur les interlocuteurs eux-mêmes. D'une part dans le cas des situations d'apprentissage qu'il s'agisse des apprentissages directs ou des apprentissages ou incidents, on s'attend à ce que l'interaction permette à l'interlocuteur le moins compétent d'acquisition de nouveaux savoirs ou savoir faire, mais on notera également que le fait même d'être confronté à un autre peut avoir un impact sur l'ensemble des participants.



De ce point de vue, l'ensemble des actions insistent sur le fait que non seulement l'interaction elle-même est un processus dynamique, qui présente des temporalités propres, avec des moments qui font avancer le dialogue, d'autres qui le bloquent, le font revenir sur des états antérieurs, affinent ou généralisent, .... Elles prennent également en compte le fait que les compétences qui se construisent dans l'interlocution répondent elles-mêmes à une dynamique propre dont la temporalité n'est pas nécessairement synchrone à celle de l'interaction verbale en tant qu'échange de mots produit temporellement mais que la construction des connaissances répond à une activité du sujet qui échappe à l'observation directe. Activité que le chercheur doit reconstruire/interpréter dans le cadre de modèles explicatifs spécifiques relevant des champs de la psychologie cognitive ou encore de la didactique des disciplines autant que dans celui de l'analyse des interactions verbales. Trois actions seront principalement conduites explorant les apprentissages médiés par ordinateur, ceux qui apparaissent dans le cadre scolaire et d'autres qui ont le plus souvent pour cadre les interactions familiales ou ordinaires.

## **2.a Actions**

### **1) Interactions médiées**

Le concept de la médiation, dans les recherches contemporaines sur l'activité humaine, trouve son origine dans les travaux princeps de L. Vygotski (1934), pour lequel la culture — avec les « outils » symboliques (le langage) et physiques (des artefacts) qu'elle comporte — structure le développement humain, s'intégrant dans la relation entre sujet et objet. Dans cette lignée, les recherches se sont centrées notamment sur trois types de médiation : (1) la médiation langagière (par ex., Hudelot, 1999) ; (2) « l'étayage » indirect de l'activité des sujets (Bruner, 1981) par une personne plus compétente dans la réalisation d'une tâche (dont l'enseignant) ; et (3) la médiation de l'artefact informatique, qui peut devenir un véritable « instrument » pour le sujet qui élabore ses propres schémas d'action (Rabardel, 1995).

Cette recherche s'inscrit dans la poursuite et la continuation des travaux du projet européen « LEAD » (« Technology-enhanced learning and problem-solving discussions : networked learning environments in the classroom », dir. M. Baker). Dans ce cadre, nous visons l'étude simultanée des trois types de médiation mentionnées ci-dessus (langage—tutelle—artefact), dans des situations complexes où des élèves du secondaire travaillent en petits groupes, en co-présence, avec des outils d'interaction médiatisée par ordinateur (réseau Intranet), avec le guidage de leur enseignant. L'enjeu est de comprendre les processus d'élaboration d'instruments et de connaissances en jeu, avec leurs interrelations, dans les interactions parlées et écrites (sur ordinateur) entre élèves, étayées par l'enseignant. Nous nous centrons sur l'étude des activités argumentatives, dont les débats à finalité pédagogique, l'élaboration de schémas argumentatifs et l'écriture collective de textes. La recherche se décline en trois étapes :

1. recueil et constitution d'un corpus d'étude, intégrant des données vidéo, des transcriptions des interactions parlées et les enregistrements automatiques des interactions médiatisées par ordinateur ;
2. élaboration d'une méthode d'analyse unifiée des différents types de médiations, en relation avec les processus d'élaboration de connaissances chez les élèves ;
3. validation et mise en œuvre de la méthode sur plusieurs corpus.

Les résultats escomptés, portant sur les processus d'élaboration de connaissances en jeu dans des activités collectives médiées, contribueront aux recherches sur les interactions langagières

menées en sciences du langage et en psychologie, et également aux fondements de la conception des situations éducatives.

## **2) Impacts de l'asymétrie des interlocuteurs sur les productions langagières du jeune enfant**

Nous nous proposons de revenir sur la notion d'étayage langagier telle qu'elle a pu être développée à la suite des travaux de Jérôme Bruner sur l'interaction de tutelle (1976), pour rendre compte de certaines conduites verbales de l'enfant lors de situations interlocutives asymétriques.

Parmi les conduites langagières pour lesquelles cette asymétrie peut avoir un impact incident, on se propose de distinguer au moins trois niveaux de fonctionnement du langage. (a) produire des énoncés au niveau phrastique. On rappellera brièvement le rôle de l'interaction comme cadre contextuel et interpersonnel dans lequel l'enfant apprend à parler en allant comme le souligne Deleau de l'interaction à la communication et de la communication au langage (Bruner, Deleau, Veneziano). (b) A un niveau textuel ou discursif, enchaîner des énoncés. Dans le cadre de la communication entre adulte et enfant, on peut en également noter comment l'enfant va s'appuyer sur les interventions d'un interlocuteur globalement plus compétent voire simplement différent pour produire à plusieurs des discours qu'il n'aurait pu produire seul, (François ; Vasseur & Hudelot). (c) Intégrer le dialogue avec autrui. On peut enfin considérer l'impact du mode d'intervention de l'adulte sur les capacités discursives manifestées par l'enfant, dans le cadre d'activité d'étayage comme procédure visant à permettre à l'enfant d'améliorer ses performances dans le cas d'une tâche complexe.

Un des objectifs de l'action sera de préciser quels sont les aspects des productions langagières qui se modifient dans l'interaction, d'indiquer les moments où cela se produit au cours du développement langagier et cognitif de l'enfant et de proposer des modèles explicatifs rendant compte des processus en cause.

### **2.b. Résultats attendus**

Les résultats issus de ces actions de recherche porteront sur les processus de transformation des productions langagières interactives, avec les notions qu'elles véhiculent, et les rôles des médiations humaines et artefactuelles dans ces transformations. Outre la publication des résultats dans des revues de recherche internationales, différentes communications (colloques et conférences invitées) sont programmées. L'opération dans son ensemble devrait donner lieu, à mi-parcours à l'organisation de journées d'études consacrées d'une part à ce que c'est qu'apprendre dans l'interaction et d'autre part à ce que c'est que rendre compte des interactions sur le long terme, celui de la rencontre, de la périodicité des rencontres, ou celui de la stabilisation des connaissances à long terme.