



HAL
open science

LATTICE - Langues, Textes, Traitements informatiques, Cognition

Rapport Hcéres

► **To cite this version:**

Rapport d'évaluation d'une entité de recherche. LATTICE - Langues, Textes, Traitements informatiques, Cognition. 2009, École normale supérieure - ENS. hceres-02031972

HAL Id: hceres-02031972

<https://hal-hceres.archives-ouvertes.fr/hceres-02031972>

Submitted on 20 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



agence d'évaluation de la recherche
et de l'enseignement supérieur

Section des Unités de recherche

Rapport d'évaluation

Unité de recherche :

Langues, Textes, Traitements Informatiques,
Cognition (LATTICE) - UMR 8094

de l'Ecole Normale Supérieure



Février 2009



agence d'évaluation de la recherche
et de l'enseignement supérieur

Section des Unités de recherche

Rapport d'évaluation

Unité de recherche

Langues, Textes, Traitements Informatiques,
Cognition (LATTICE) - UMR 8094

de l'Ecole Normale Supérieure



Le Président
de l'AERES

Jean-François Dhainaut

Section des unités
de recherche

Le Directeur

Pierre Glorieux

Février 2009



Rapport d'évaluation



L'Unité de recherche :

Nom de l'unité : Langues, Textes, Traitements Informatiques, Cognition (LATTICE)

Label demandé : UMR

N° si renouvellement : 8094

Nom du directeur : M. Michel CHAROLLES

Université ou école principale :

ENS

Autres établissements et organismes de rattachement :

Université Paris 3

Date(s) de la visite :

6 février 2009



Membres du comité d'évaluation

Président :

M. Wolfgang WILDGEN, Université de Brême, Allemagne

Experts :

M. Bruno BACHIMONT, Université de technologie de Compiègne

Mme Maya HICKMANN, Université Paris 8

Mme Adeline NAZARENKO, Université Paris 13

Mme Jeannine RICHARD-ZAPPELLA, Université d'Amiens

Expert(s) représentant des comités d'évaluation des personnels (CNU, CoNRS, CSS INSERM, représentant INRA, INRIA, IRD...) :

M. Thierry OLIVE, représentant le CNRS

M. Henri PORTINE, représentant le CNU

Observateurs

Délégué scientifique de l'AERES :

Mme Sandra LAUGIER

Représentant de l'université ou école, établissement principal :

M. Jean-Claude DARMON, ENS

Représentant(s) des organismes tutelles de l'unité :

M. Cuong LE VAN, DSA CNRS

M. Pierre CIVIL, VP recherche, Université Paris 3



Rapport d'évaluation

1 • Présentation succincte de l'unité

- effectif : 13 dont 7 enseignants-chercheurs (3 PR, 4 MCF) et 6 chercheurs (2 DR, 4 CR)
- ingénieurs : 2 (un IE, un IR)
- doctorants : 18, 2 Post doctorants
- techniciens et administratifs : 1
- nombre de HDR : 7, nombre de HDR encadrant des thèses id.
- nombre de thèses soutenues et durée moyenne lors des 4 dernières années : 4/an, durée moyenne 4 ans
- nombre de membres bénéficiant d'une PEDR 2
- nombre de publiants 13, 100% de publiants

2 • Déroulement de l'évaluation

La visite s'est déroulée sur une après midi en présence des représentants des tutelles.

Elle s'est déroulée en trois phases : un exposé général par le directeur présentant très clairement les principales orientations du laboratoire et un rappel de son historique, suivi d'un bilan des opérations et des projets par le directeur-adjoint et les membres responsables de chacun des axes. Puis des présentations des programmes en cours : synthétiques, brillantes, majoritairement présentées par les jeunes chercheurs de l'équipe, ces interventions ont été fort appréciées en donnant une vision vivante, dynamique et interactive des travaux et perspectives de l'unité. Enfin des entretiens spécifiques avec les chercheurs et les enseignants chercheurs, les doctorants et les ITA ont permis au comité de mieux cerner le mode de travail de l'unité.

La visite du laboratoire a ainsi permis de compléter utilement le document écrit (qui était un peu lacunaire) notamment en mettant en valeur des questionnements transversaux aux axes de recherches. Elle a mis en évidence la vitalité des équipes, la cohésion entre les membres du laboratoire et le soutien apporté de façon unanime au projet du laboratoire et à l'équipe de direction qui le porte.

Elle a permis enfin de mieux mettre en évidence la cohérence du projet scientifique de l'équipe et son originalité.

La rencontre avec les doctorants a montré le soutien important, tant scientifique que logistique, qui leur est apporté, ainsi que leur vive appréciation de l'ambiance intellectuelle de l'unité, qui constitue incontestablement un milieu intellectuel riche et stimulant.

3 • Analyse globale de l'unité, de son évolution et de son positionnement local, régional et européen

Le LATTICE (Langues, Textes, Traitements Informatiques, Cognition) résulte de la fusion en janvier 2000 de deux équipes, l'EP LTM "Langues, Textes, Modèles" (autonomisation de l'antenne parisienne du laboratoire ELSAP de Caen) et de TALANA, EA de Paris 7.

L'équipe a connu des transformations importantes dans la période récente. La directrice de l'unité désignée pour le quadriennal 2006-2009 ayant demandé en 2008 la reconnaissance d'une nouvelle unité, elle a quitté l'équipe. Le changement n'a nui ni à la cohérence de l'unité ni à son développement. Il a permis un



renouvellement du projet et un resserrement de l'unité, très unie et motivée autour de sa nouvelle direction. Le nouveau directeur et le directeur adjoint font l'unanimité et il règne manifestement dans l'équipe une ambiance chaleureuse et un dynamisme impressionnant. L'attractivité nouvelle de l'unité est marquée par le recrutement de jeunes chercheurs de très bon niveau et l'augmentation forte du nombre de doctorants.

Le nouveau projet implique une légère modification dans les tutelles : la tutelle principale reste l'ENS Ulm, la seconde passant de l'université Paris 7 (dont la présence dans l'unité est résiduelle) à l'université Paris 3, très impliquée. Le LATTICE est implanté essentiellement sur un seul site, celui de Montrouge dans des locaux de l'ENS Ulm, mais des locaux dans une prochaine maison de la recherche de l'université Paris 3 sont envisagés.

Le rapport scientifique des années 2006-2009 a une structure claire qui montre la portée des recherches effectuées. Elles couvrent d'une part beaucoup de problèmes actuels en linguistique : de la phrase au discours, tout en accentuant des points critiques (ou problématiques), tels que la prédication averbale/verbale et la connexion syntaxique dans le domaine de la phrase (subordination et comparaison) et les relations / connecteurs dans le discours à caractère adverbial (par exemple les adverbes en -ment).

Un point fort concerne l'étude des repérages spatiaux et temporels.

Au-delà de la grammaire au sens strict, deux volets sur l'interdisciplinarité sont ouverts : le traitement automatique des textes et les aspects cognitifs de la grammaire.

Si les analyses concrètes concernent surtout le français, d'autres langues indo-européennes et non-indoeuropéennes (fino-ougriques et japonais) sont considérées et l'aspect typologique est pris en compte.

La méthodologie employée met l'accent sur l'analyse qualitative et quantitative des corpus et les outils informatiques nécessaires sont développés. Mais l'expérimentation psycholinguistique est également présente.

Ces recherches sont très pertinentes pour le développement actuel de la linguistique et très originales.

Les chercheurs et les enseignants-chercheurs qui composent les différents axes ont une solide réputation nationale et internationale pour la qualité de leurs travaux, de leurs publications, ainsi que pour la production d'outils, et notamment de logiciels.

L'équipe possède une activité scientifique importante, et devrait à ce titre mieux se positionner sur le plan national et international. De même, les options scientifiques prises par l'équipe devraient être davantage explicitées et situées par rapport aux discussions du domaine et aux courants qui le traversent.

L'unité a évolué de manière très positive dans des circonstances difficiles et a pu se reconstruire autour d'un projet scientifique resserré, fort des travaux antérieurs des membres restant dans l'unité et scientifiquement cohérent. Au plan institutionnel et pédagogique, le LATTICE paraît déjà bien intégré à l'Université Paris 3 et à son Ecole Doctorale. On regrette que les particularités de l'ENS rendent l'intégration de l'unité au sein de cette école un peu plus difficile alors que le LATTICE peut offrir des perspectives de recherche en linguistique très intéressantes pour les élèves, littéraires et philosophes, de l'ENS. Cela suppose une politique plus active de la part des chercheurs auprès de la direction de l'ENS et de ses élèves, mais également une prise de contact facilitée entre le LATTICE et les élèves arrivant à l'ENS, comme l'a d'ailleurs souligné le représentant de l'Ecole.

4 • Analyse équipe par équipe et par projet

Les recherches menées vont de la description linguistique (marqueurs grammaticaux, syntagmes et prédications, structuration du discours) à la modélisation et au traitement automatique, en passant par des réflexions sur la cognition et des expérimentations psycholinguistiques.

L'équipe poursuit un certain nombre de programmes explorés précédemment : la comparaison, la prédication averbale, les aspects spatiaux. Elle les développe et les complète en considérant : la grammaticalisation (dynamique historique de la grammaire), l'interaction et les transitions entre unités prédictives verbales et averbales. Deux volets nouveaux sont ouverts : les éléments initiaux et leur relation au sujet et les mécanismes de référence pronominale (y inclus leur ambiguïté).



L'équipe paraît particulièrement à même de réaliser ces projets grâce à la continuité de ses programmes et à leur expansion cohérente.

On note aussi une ouverture théorique qui considère les développements internationaux dans le domaine de la dynamique des langues (par exemple les théories de la grammaticalisation, des chaînes grammaticales etc.) et de la grammaire cognitive (avec une dimension fonctionnelle ajoutée).

Par le biais de collaborations fructueuses (Paris 8 et Paris 5), les chercheurs ont su intégrer des méthodologies relevant de la psycholinguistique aux questionnements émergeant des recherches en linguistique - la démarche amorcée consistant à mettre à l'épreuve certaines de leurs hypothèses dans une interface interdisciplinaire langage-cognition.

Simultanément, les recherches en linguistique diachronique ont été renforcées dans l'ensemble du dispositif par le recrutement d'un excellent jeune chercheur CNRS et l'obtention d'un projet franco-allemand.

Le travail instrumenté sur corpus apparaît comme un commun dénominateur des travaux de recherche du LATTICE - comme c'est souvent le cas aujourd'hui dans les recherches en linguistique. La restructuration du LATTICE a permis de redéfinir le rôle, dans l'unité, du traitement automatique des langues, qui n'est plus présenté comme un axe transversal et mais comme un choix méthodologique.

Ce nouveau positionnement est judicieux. L'effort fourni dans le développement d'outils d'observation en corpus doit être mené en concertation avec la communauté nationale et internationale : en effet le LATTICE a une originalité à faire valoir dans l'analyse comparative (en diachronie ou en synchronie) et dans l'analyse détaillée du fonctionnement des marqueurs linguistiques.

Certaines activités sont particulièrement innovantes, comme les outils d'instrumentation des corpus : il s'agit certes d'une tendance désormais commune, mais le LATTICE apporte une note originale qu'il devrait assumer explicitement et mieux diffuser :

- le rôle des outils dans la construction du savoir du domaine (volet épistémologique)
- instrumentation graphique et interactive des corpus (volet méthodologique).

5 • Analyse de la vie de l'unité

La gestion du laboratoire semble excellente. Il s'agit d'une équipe très structurée, très organisée ; cette bonne gouvernance apparaît notamment dans le soutien apporté à l'encadrement des doctorants et des jeunes chercheurs, qui l'ont fortement souligné lors des entretiens.

L'équipe est petite mais très soudée et elle sait avoir un management souple et réactif. Elle a su attirer et retenir ces dernières années de jeunes chercheurs, indiquant ainsi la pertinence et l'actualité de ses thèmes scientifiques de même que la qualité de l'environnement de travail qu'elle offre. L'ambiance apparaît très propice aux échanges scientifiques et l'ensemble de l'unité est géré de manière efficace, grâce aussi à un personnel motivé.

On note cependant un taux encore insuffisant d'HDR parmi les membres permanents, ce qui entraîne un déséquilibre dans l'encadrement des thèses du laboratoire. Il faut encourager les membres de l'équipe qui ont la possibilité de le faire à préparer leur HDR, afin d'obtenir une organisation plus équilibrée de l'animation et de l'encadrement scientifique.

En particulier, l'équipe comprend des membres seniors d'envergure scientifique de premier plan dont le futur départ à la retraite doit être planifié et géré.

L'équipe comprend un nombre important de doctorants. Si on note avec intérêt que certains obtiennent un financement CIFRE auprès des entreprises, la majorité n'est pas financée. Pratique usuelle en SHS mais regrettable dans ces disciplines : on ne peut qu'encourager le laboratoire à poursuivre ses efforts pour proposer des financements à ses doctorants, notamment auprès de la tutelle universitaire.

Enfin, le lien avec la tutelle principale semble insuffisant et on attend de l'ENS qu'elle intègre plus explicitement le LATTICE, équipe d'excellence dont les recherches originales sont au croisement de ses



disciplines prioritaires (philosophie, littérature, informatique, sciences cognitives), dans sa politique scientifique, afin de mieux définir dans le quadriennal les modalités de son soutien institutionnel et matériel.

6 • Conclusions

Points forts

- L'originalité et la finesse des recherches : la plupart des opérations ont recours à des analyses fines sur de larges corpus, qui ont impliqué le développement de méthodes innovantes (logiciels, systèmes d'annotations et de recherche). La capacité à proposer un nouvel usage de la linguistique de corpus, le travail sur différentes langues (germaniques mais aussi finnoise et hongroise), la richesse de l'approche méthodologique, qui permet d'occuper une position originale dans le domaine. Les corpus mis à la disposition de la communauté scientifique et les logiciels d'analyse prosodique (Analor) et sémantique (Anasem) ont une valeur pertinente en France et à l'international.
- L'interdisciplinarité : par le biais de collaborations fructueuses (Paris 8 et Paris 5), les chercheurs ont su intégrer des méthodologies relevant de la psycholinguistique aux questionnements émergents des recherches en linguistique - la démarche amorcée consistant à mettre à l'épreuve certaines des hypothèses dans une interface interdisciplinaire langage-cognition. La pluralité d'approche des textes littéraires est un élément d'attractivité pour les étudiants.
- L'environnement scientifique, notamment les liens étroits avec l'Ecole Doctorale de l'université Paris 3 et la localisation à l'ENS. La relation entre recherche fondamentale et recherche appliquée est bien équilibrée.
- Le LATTICE est impliqué dans plusieurs projets ANR et européens et leader dans deux programmes.
- Malgré les difficultés rencontrées ces dernières années, le bilan du laboratoire témoigne d'un niveau de publication important de chacun des membres du laboratoire et d'un encadrement de doctorants satisfaisant (nombreuses soutenances de thèses) . On constate également un nombre important de publications impliquant plusieurs auteurs du laboratoire, et ce dans tous les axes de recherche.
- L'ambiance de travail qui règne dans le laboratoire est particulièrement chaleureuse et constructive. La direction et la co-direction ont été unanimement choisies et sont fort appréciées. Le laboratoire est aussi clairement ouvert sur l'extérieur, avec de nombreuses collaborations scientifiques avec d'autres chercheurs et équipes nationales et internationales.

Points faibles

- Il est à regretter que les publications en langue anglaise ne soient pas plus nombreuses, qu'elles ne soient le fait que de quelques collègues. Il est recommandé au directeur de l'unité de veiller à augmenter le nombre de publications en anglais de chacun des membres du laboratoire, y compris des doctorants.
- L'équipe souffre d'un déficit d'image ; il faut mieux situer les options de l'équipe dans les débats actuels du domaine et se positionner dans la cartographie nationale et internationale de la recherche du domaine.
- La tutelle ENS est insuffisamment impliquée au plan scientifique comme pédagogique.
- On peut avoir quelque inquiétude en ce qui concerne la pyramide des âges. Un rajeunissement des cadres est souhaité à court terme afin de permettre au laboratoire de continuer son développement.



Recommandations

- Il est important que les encadrements de thèses soient mieux répartis entre les différents membres du laboratoire, mais aussi que soit associés à ces encadrements des personnels qui n'ont pas encore obtenu leur HDR (par exemple par des co-directions) afin de préparer l'évolution de leur carrière.
- Il faut poursuivre l'effort d'articulation des différentes opérations afin de mettre en route un vrai projet collectif énoncé comme tel, et pas seulement marqué par la récurrence des noms des chercheurs et des enseignants-chercheurs dans l'énoncé des opérations de recherche.
- Il faut agir auprès des tutelles (afin de les impliquer davantage) afin d'obtenir un réel soutien scientifique et leur implication dans les recherches de l'unité ; notamment amener l'ENS à mieux inscrire la linguistique dans ses formations. Les responsables sont invités à définir en concertation avec la direction de l'ENS une politique active pour faire découvrir la recherche en linguistique aux élèves littéraires de l'ENS.
- Le positionnement de l'unité par rapport à l'environnement national et international des recherches en linguistique doit être clarifié afin de mettre en valeur la spécificité de ses recherches.
- Enfin, il est recommandé à l'unité de mieux positionner les collaborations avec la psycholinguistique dans le projet proposé. Plus précisément, s'agit-il à l'avenir de simplement continuer ces collaborations, ou s'agit-il de l'amorce d'une ouverture du laboratoire à la psycholinguistique en vue de recrutements ultérieurs de chercheurs et enseignants-chercheurs dans ces disciplines ?
- Il faut tenter de compenser le « trou » qui apparaît dans la pyramide des âges (l'équipe s'appuie sur de très jeunes chercheurs et des chercheurs très confirmés), en recrutant des chercheurs d'âge intermédiaire (CR1), et au moins à anticiper les départs, dans les années qui viennent, en préparant les jeunes chercheurs à prendre la relève.

Note de l'unité	Qualité scientifique et production	Rayonnement et attractivité, intégration dans l'environnement	Stratégie, gouvernance et vie du laboratoire	Appréciation du projet
A+	A+	A	A+	A



Lattice

UMR 8094



Langues, Textes, Traitements Informatiques, Cognition

Paris le 26/03/09

Ecole Normale Supérieure (Ulm)
1 rue Maurice Arnoux 92120 Montrouge

Michel Charolles

Directeur

Réactions au pré-rapport établi par le Comité de Visite de l'AERES

Nous remercions vivement les membres du Comité pour les appréciations qu'ils portent sur le bilan et le projet que nous avons soumis et pour le profit que nous ne manquerons pas d'en retirer dans les travaux que nous menons. Les points forts relevés nous confortent dans les directions que nous avons prises. Les points faibles et les recommandations mettent le doigt sur des aspects qui, pour l'essentiel, ne nous avaient pas échappé (cf. l'autoévaluation remise le jour de la visite).

Les informations factuelles mentionnées dans le rapport n'appellent aucune observation de notre part.

Concernant les points faibles et les recommandations mentionnées dans le rapport, nous sommes bien entendu tout à fait disposés à multiplier les soumissions en langue anglaise dans des revues internationales. Il faut bien voir cependant que certaines études que nous menons portent sur des formes d'expression relativement spécifiques à notre langue et dont l'apport ne peut guère être apprécié et évalué que par des spécialistes (français ou étrangers). La publication des résultats dans des revues en français nous permet de bénéficier des réactions de ces spécialistes en premier lieu. Mais, lorsqu'il s'avère que ces résultats et les perspectives qui en découlent sont généralisables et peuvent contribuer à la discussion de questions théoriques débattues dans la communauté internationale, nous sommes naturellement prêts à les soumettre à des revues de plus large diffusion, ce que nous n'avons du reste pas manqué de faire durant le précédent quadriennal. Nous nous attacherons à le faire encore plus nettement durant le prochain quadriennal, en mettant en place un dispositif d'aide aux chercheurs, enseignants-chercheurs et doctorants de l'unité qui le souhaitent.

Pour ce qui est du "déficit d'image" de notre équipe relevé par le Comité, notre positionnement parmi les grands courants de recherche en linguistique n'est pas toujours évident dans la mesure où nous pratiquons des analyses résolument empiriques, fondées sur des corpus aussi larges et représentatifs que possible, en synchronie et en diachronie, sur le

français écrit et parlé, et où les généralisations sont souvent difficiles. Cela dit, les travaux que nous menons se rattachent à des courants que nous connaissons bien et nous entretenons des liens très étroits à l'échelle internationale avec des chercheurs et des équipes qui travaillent dans le même esprit que nous. Ces liens peuvent être explicités, nous le ferons en mettant l'accent, dans nos contributions, sur leur enjeu et leur apport théorique général.

Le souci d'ajouter aux données sur corpus des données comportementales recueillies par le biais d'expérimentations psycholinguistiques s'explique, de la même façon, par la volonté de coller aux faits empiriques. Le recours à certaines méthodologies expérimentales est maintenant à notre portée, nous souhaitons développer ce secteur après les essais concluants que nous avons pu faire. Dans la demande de moyens adressée au CNRS pour 2009, nous avons demandé un CDD de 2 ans en psycholinguistique afin d'avancer dans ce secteur où notre expérience est encore limitée. Notre objectif à court terme est d'installer dans le laboratoire un secteur psycholinguistique et nous demanderons un poste de CR durant le quadriennal. Nous visons une certaine autonomie dans ce domaine, tout en continuant à renforcer nos liens avec les laboratoires spécialisés. Les collaborations que nous avons déjà nouées et celles que nous souhaitons développer dans ce secteur sont essentielles et participent d'un souci d'interdisciplinarité qui est vital pour notre équipe. Ce souci vaut en particulier pour les études sur le français parlé, sur la référence et pour le développement d'outils informatiques destinés au traitement de données linguistiques, domaines où nous avons déjà des collaborations bien établies que nous chercherons, là aussi, à étoffer.

Concernant les liens avec l'ENS, le nouveau séminaire de "Linguistique et analyse des textes littéraires" qui vient de commencer au second semestre 2008-2009 a trouvé sa place, avec une dizaine d'élèves de l'École. Nous ferons le bilan avec les responsables du LILA, mais l'expérience semble concluante et nous souhaitons la reconduire l'année prochaine. Plusieurs membres de notre équipe vont participer à un groupe de travail mis en place par A-M. Paillet. Nous allons par ailleurs reprendre, à partir de la rentrée universitaire prochaine, le cycle de conférences (que nous avons instauré lors du précédent quadriennal) en invitant à l'École des linguistes de grande réputation et nous allons organiser à l'ENS, fin septembre, un colloque international "Linguistic and Psycholinguistic Approaches to Text Structuring". Nous prendrons durant le prochain quadriennal toutes les initiatives et répondrons à toutes les demandes de nature à renforcer notre positionnement à l'ENS et pour participer à la formation des élèves. Nous ferons de même à Paris III où la majorité des enseignants-chercheurs de notre équipe sont en poste et où plusieurs chercheurs sont déjà impliqués.

Enfin, pour ce qui est de la pyramide des âges et de l'anticipation des départs à la retraite, notre équipe compte

- une MC déjà habilitée ainsi qu'une MC qui achève son HDR, qui devraient obtenir un poste de professeur en cours de quadriennal,
- trois CR en position de passer leur HDR à bref délai.

Une candidate sur un poste de CR1 et de CR2 a demandé notre laboratoire cette année et, notre UMR ayant su, durant le précédent quadriennal, attirer deux chercheurs, il n'y a pas de raison pour que cette dynamique s'arrête, surtout si on nous en donne les moyens, comme nous ne manquerons pas de le demander (cf. ci-avant). Nous ferons également des démarches auprès de l'ENS et de Paris III pour que notre UMR bénéficie de recrutements sur des postes vacants ou nouveaux durant les campagnes du quadriennal.

Michel Charolles