



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. 2013, École normale supérieure - ENS, Centre national de la recherche scientifique - CNRS, Université Sorbonne Nouvelle - Paris 3. hceres-02031971

HAL Id: hceres-02031971

<https://hal-hceres.archives-ouvertes.fr/hceres-02031971v1>

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

Évaluation de l'AERES sur l'unité :
Langues, Textes, Traitements Informatiques, Cognition
LATTICE

sous tutelle des

établissements et organismes :

Université Sorbonne Nouvelle - Paris 3

École Normale Supérieure

Centre National de la Recherche Scientifique



Décembre 2012



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

Section des Unités de recherche

Le Président de l'AERES

Didier Houssin

Section des Unités
de recherche

Le Directeur

Pierre Glaudes



Notation

À l'issue des visites de la campagne d'évaluation 2012-2013, les présidents des comités d'experts, réunis par groupes disciplinaires, ont procédé à la notation des unités de recherche relevant de leur groupe (et, le cas échéant, des équipes internes de ces unités). Cette notation (A+, A, B, C) a porté sur chacun des six critères définis par l'AERES.

NN (non noté) associé à un critère indique que celui-ci est sans objet pour le cas particulier de cette unité ou de cette équipe.

Critère 1 - C1 : Production et qualité scientifiques ;

Critère 2 - C2 : Rayonnement et attractivité académique ;

Critère 3 - C3 : Interaction avec l'environnement social, économique et culturel ;

Critère 4 - C4 : Organisation et vie de l'unité (ou de l'équipe) ;

Critère 5 - C5 : Implication dans la formation par la recherche ;

Critère 6 - C6 : Stratégie et projet à cinq ans.

Dans le cadre de cette notation, l'unité de recherche concernée par ce rapport a obtenu les notes suivantes :

- Notation de l'unité : **LANGUES, TEXTES, TRAITEMENTS INFORMATIQUES, COGNITION**

C1	C2	C3	C4	C5	C6
A+	A	B	A+	A	A



Rapport d'évaluation

Nom de l'unité :	Langues, Textes, Traitements informatiques, Cognition
Acronyme de l'unité :	LATTICE
Label demandé :	UMR CNRS
N° actuel :	8094
Nom du directeur (2012-2013) :	M. Thierry POIBEAU
Nom du porteur de projet (2014-2018) :	M. Thierry POIBEAU

Membres du comité d'experts

Président :	M ^{me} Annie KUYUMCUYAN, Nancy
Experts :	M ^{me} Mireille BILGER, Perpignan (Représentante du CNU)
	M ^{me} Cécile FABRE, Toulouse
	M ^{me} Claire GARDENT, Nancy (Représentante du CoNRS)
	M ^{me} Corinne KALFA, Nice (Représentante du CoNRS/ITA)
	M. Michel PIERRARD, Bruxelles (Belgique)

Délégué scientifique représentant de l'AERES :

M^{me} Catherine SCHNEDECKER

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

M. Carle BONAFOUS-MURAT, VP Conseil Scientifique Paris 3

M. Guillaume BONNET, Directeur Lettres, ENS

M. Stefano BOSI, DAS INSHS, CNRS



1 • Introduction

Historique et localisation géographique de l'unité

Le LATTICE est une UMR sous la triple tutelle du CNRS, de l'École normale supérieure et de l'Université Paris 3 Sorbonne nouvelle. Le laboratoire a été créé en janvier 2000 et est localisé dans les locaux de l'École normale supérieure à Montrouge. Aucun changement dans les tutelles n'est intervenu depuis la dernière évaluation de l'AERES en 2009.

Équipe de Direction

Le laboratoire est placé sous la direction de M. Thierry POIBEAU (directeur) et de M^{me} Sophie PREVOST (directrice adjointe). Cette nouvelle direction a pris ses fonctions au cours de l'année 2012 suite à un vote en assemblée générale du laboratoire en janvier 2012 et une validation de l'ensemble des tutelles en début d'année.

Jusqu'à la fin 2011, la direction du laboratoire était assurée par M. Michel CHAROLLES (directeur) et M. Bernard VICTORRI (directeur adjoint).

Effectifs de l'unité

Effectifs de l'unité	Nombre au 30/06/2012	Nombre au 01/01/2014	2014-2018 Nombre de produisants du projet
N1 : Enseignants-chercheurs titulaires et assimilés	8	7	7
N2 : Chercheurs des EPST ou EPIC titulaires et assimilés	5	5	5
N3 : Autres personnels titulaires (n'ayant pas d'obligation de recherche)	5 (4,3 ETP)	4 (4 ETP)	2
N4 : Autres enseignants-chercheurs (PREM, ECC, etc.)	1	2	2
N5 : Autres chercheurs des EPST ou EPIC (DREM, Post-doctorants, visiteurs etc.)	1	1	1
N6 : Autres personnels contractuels (n'ayant pas d'obligation de recherche)	0	0	0
TOTAL N1 à N6	20	19	17
Taux de producteurs	100 %		



Effectifs de l'unité	Nombre au 30/06/2012	Nombre au 01/01/2014
Doctorants	14	
Thèses soutenues	12	
Post-doctorants ayant passé au moins 12 mois dans l'unité *	3	
Nombre d'HDR soutenues	4	
Personnes habilitées à diriger des recherches ou assimilées	10	12



2 • Appréciation sur l'unité

Points forts et possibilités liées au contexte

Les points forts du LATTICE (Langues, Textes, Traitements Informatiques, Cognition) sont sa grande visibilité en France et la qualité incontestable de sa production scientifique, dans des secteurs variés de la linguistique - en particulier le discours, la diachronie, le lexique, et désormais le traitement automatique des langues (TAL). On relève également l'importance de ses partenariats scientifiques qui ont accru son rayonnement, y compris sur le plan international, et lui fournissent des financements extérieurs non négligeables. C'est également un laboratoire qui se caractérise par l'ouverture à l'innovation technologique et une capacité importante de renouvellement méthodologique. Un des atouts de l'unité consiste en effet à combiner une démarche de linguistique de corpus et une démarche expérimentale de nature psycholinguistique pour croiser les types de données et les modes de validation. Une évolution convaincante se dessine dans la volonté de s'ouvrir à la diversité des langues sur la base d'une forte tradition de centrage sur l'étude du français.

Points à améliorer et risques liés au contexte

Le départ à la retraite de certaines de ses figures les plus marquantes constitue une étape importante dans la vie de l'unité. Si tout a été fait pour assurer la transition entre l'ancienne et l'actuelle équipe de direction, les nouvelles orientations de recherche pèchent parfois par une certaine abstraction ou une trop grande généralité. Il y aurait également lieu de veiller à la cohérence thématique des différents travaux menés dans l'unité et à leur articulation, qui doivent être plus clairement affichées.

Au plan institutionnel et pédagogique, compte tenu de ses particularités comme des récents mouvements de personnel, les enseignants-chercheurs notamment, l'unité court le risque de continuer à occuper une place relativement marginale vis-à-vis de ses deux tutelles, l'université Paris 3 d'une part, l'ENS de l'autre. Les liens avec ces deux institutions doivent être renforcés - notamment au travers de leurs Écoles Doctorales respectives - ainsi que les représentants de ces tutelles en ont eux-mêmes exprimé la volonté, de sorte que les étudiants soient mieux informés de l'existence et des activités du LATTICE et des champs de recherche explorés par l'unité. On relève un certain nombre de tentatives innovantes en ce sens, qui sont à poursuivre et à développer.

Recommandations

Hormis le renforcement des liens avec les tutelles mentionné ci-dessus, pour une meilleure intégration de l'unité dans le tissu universitaire de proximité, le LATTICE aurait également intérêt à rendre son unité plus explicite, aussi bien quant aux thématiques de recherche qu'au plan des méthodes et des options scientifiques, et ce pour mieux mettre en valeur ses grandes potentialités.



3 • Appréciations détaillées

Appréciation sur la production et la qualité scientifiques

La qualité scientifique de la production de l'unité est de premier ordre, dans les différents domaines où elle se situe. On relèvera par exemple le développement et la mise en ligne du logiciel d'analyse prosodique du discours ANALOR, du logiciel d'analyse de l'écrit ANALEC, le développement et la mise en ligne de la base Structures Comparatives du français. Imputés dans le bilan à l'opération transversale « Linguistique outillée », ces résultats s'appuient en réalité sur une synergie des recherches menées dans l'unité, et témoignent donc en sus d'une parfaite intégration entre la recherche théorique et sa mise en œuvre au plan informatique. L'impact de ces travaux est réel au sein du monde académique, par les discussions et les débats qu'il suscite, et les avancées théoriques et méthodologiques auxquelles il contribue. L'unité est par ailleurs soucieuse d'ouverture par le biais de publications dans les revues nationales et étrangères. On relève une proportion satisfaisante de publications dans des revues à comité de lecture, ainsi qu'un pourcentage très honorable de publications en anglais.

Appréciation sur le rayonnement et l'attractivité académiques

L'unité est partie prenante de nombreux projets nationaux et internationaux. Elle participe à deux labex (TransferS et EFL) ainsi qu'à différents projets ANR y compris en tant que laboratoire porteur - projets SFA, SRCMF, ORFEO, etc.- On relève également des participations à la Grande Grammaire du Français, un projet transversal de contribution à la Grande grammaire historique du français. Les partenariats avec les universités étrangères sont également assurés grâce aux labex et ANR précédemment évoqués ainsi qu'à d'autres collaborations européennes ou internationales. L'unité a invité des chercheurs étrangers de notoriété internationale pour des cycles de conférences.

Notons aussi l'implication de l'unité dans la création et la gestion de la revue en ligne Discours, ainsi que la présence de ses membres dans les comités de rédaction de revues importantes dans leurs domaines respectifs. En sus de leur contribution propre à la recherche, les membres de l'équipe veillent donc aussi à intervenir dans sa communication au travers de différents supports et instances de médiation, d'organisation de colloques (LPTS, sept. 2009) ou d'école thématique (en collaboration avec MODYCO en 2011).

Appréciation sur l'interaction avec l'environnement social, économique et culturel

L'interaction avec l'environnement, hormis les publications, corpus et bases de données, est notamment assurée par le pôle TAL de l'unité, qui génère « naturellement » des liens avec les industries du domaine, notamment en termes de formation (organisation de journées d'études par exemple). C'est dans ce cadre qu'ont été aussi financées deux thèses. L'équipe a également déposé cinq logiciels en open source auprès du CNRS. Elle collabore par ailleurs au mouvement des « humanités numériques » en partenariat avec le laboratoire d'archéologie de l'ENS afin de développer une véritable compétence en la matière au sein de sa tutelle principale, même si le domaine d'application (les Cartes Archéologiques de la Gaule) semble a priori plutôt éloigné de son sujet principal,

Un autre domaine d'interaction privilégié avec l'environnement est l'enseignement du français langue étrangère avec un dictionnaire des verbes du français et sa version électronique sur CD-Rom, ou la réflexion sur les langues de spécialité, l'anglais notamment.

Notons enfin des tentatives intéressantes de vulgarisation hors du monde de la recherche proprement dit : lancement de la revue Sciences/Lettres et rédaction de deux de ses numéros par des membres de l'unité, participation à des émissions radio-diffusées, à divers ouvrages de vulgarisation.

L'unité est donc partie prenante de multiples actions de promotion de la recherche et de collaboration avec des partenaires extra-académiques, ce qui lui permet d'entrer en synergie avec le tissu socio-économique et culturel et de nourrir les recherches menées au sein de l'unité par des problématiques issues des industries numériques impliquées dans la culture et les médias.



Appréciation sur l'organisation et la vie de l'unité

Le LATTICE forme une unité de format raisonnable dont le mode de gouvernance ne semble pas poser de problème grâce aux assemblées générales tenues trois fois par an. L'ancienne comme la nouvelle équipe de direction ont été unanimement appréciées.

La formation des personnels, tant techniques que de recherche, est assurée par des stages ou des manifestations diverses. On s'étonne cependant que l'unique représentante du personnel technique dans l'unité de recherche (ITA) n'ait pas suivi une évolution de carrière conforme à son investissement et à son implication personnelle dans la vie de l'unité.

Le comité d'experts ne peut que vivement encourager à une remédiation rapide de cette anomalie, seul point noir de ce chapitre au demeurant tout à fait satisfaisant par ailleurs.

Eu égard à la localisation géographique de l'unité, les locaux peuvent être considérés comme satisfaisants quant à leur taille. Le site web est régulièrement entretenu et le personnel est dûment informé de la politique scientifique de l'unité et des programmes de recherche.

Appréciation sur l'implication dans la formation par la recherche

Grâce à sa rencontre avec les doctorants lors de sa visite sur site, le comité d'experts a pu s'assurer de leur satisfaction quant à leur accompagnement et à leur encadrement dans l'unité de recherche. Un témoignage a particulièrement pointé la grande qualité de l'accueil dans l'équipe en termes de disponibilité pour l'encadrement des doctorants. Au LATTICE, tout est fait en termes d'accueil, d'écoute, de séminaires et de formation à destination des doctorants. On relève en particulier les « journées de doctorants » (trois à quatre par an), une liste de diffusion propre, le soutien à leur participation à diverses manifestations (écoles d'été, colloques, etc.)

L'École Doctorale 268 (Langage et langues : description, théorisation, transmission) de Paris 3 a été dirigée de 2007 à 2009 par le précédent directeur de l'unité, lui-même enseignant-chercheur à Paris 3. Pour Paris 3 encore, la formation PluriTAL (master de traitement automatique des langues commun à Paris 3, Paris 10 et INaLCO) est coordonnée par un autre enseignant-chercheur rattaché au LATTICE. On peut cependant craindre qu'avec le départ à la retraite de deux enseignants-chercheurs de Paris 3, par ailleurs figures éminentes du LATTICE, les liens avec cette ED, et avec Paris 3 de manière générale, ne se distendent. Dans un contexte de concurrence de plus en plus ardue entre les acteurs du monde de la recherche, sur fond de baisse du nombre des nouvelles vocations dans la recherche fondamentale ou appliquée en sciences du langage, l'unité ne doit pas céder la place qu'elle occupait jusqu'ici, grâce à ses enseignants-chercheurs, à Paris 3. Le cas de l'École Normale Supérieure et de son École Doctorale 540 est un peu différent, eu égard aux particularités de cette institution, car aucun de ses enseignants-chercheurs n'est par exemple membre du LATTICE. C'est pour remédier à cette anomalie qu'un « projet innovant » a été mis en place sur « la lecture linguistique de textes littéraires », à l'initiative de ces mêmes enseignants de Paris 3 désormais émérites. Pour que l'unité continue à tenir une place dans la formation par la recherche dispensée par ses tutelles, il serait souhaitable que de telles initiatives aient une suite et qu'en particulier les chercheurs CNRS continuent à intervenir, comme ils le font déjà, dans diverses formations universitaires nationales et internationales.

Appréciation sur la stratégie et le projet à cinq ans

En raison de la recomposition qui l'affecte suite à un renouvellement important de son équipe, l'unité est actuellement dans une phase de transition. On peut verser au crédit de l'équipe de direction antérieure d'avoir anticipé cette évolution et d'avoir essayé de l'accompagner, notamment en procédant à son propre remplacement, dans un esprit constructif et consensuel. Nul doute que la synergie humaine n'ait été de la sorte acquise, ce qui est un point important. On reste un peu plus dubitatif quant à la synergie scientifique, étant donné la variété des projets, laquelle va jusqu'à une certaine dispersion. Si celle-ci est en un sens un signe incontestable de dynamisme, elle suscite néanmoins un certain scepticisme, tant au niveau de la cohérence globale du projet qu'au sein de chacun des axes, comme l'analyse détaillée le montrera ci-après.



En somme, on peut regretter que l'élargissement disciplinaire qui se dessine de façon prometteuse du côté du TAL ne soit pas accompagné d'un engagement similaire du côté de la psycholinguistique, laquelle était pourtant une marque de fabrique de l'innovation apportée par l'équipe jusqu'ici. Certes, ce rééquilibrage risque de jouer en faveur des partenariats extra-académiques, industriels notamment, mais il ne doit cependant pas nuire à l'articulation avec la recherche fondamentale, qui doit conserver toute sa place au sein d'une unité d'aussi grande qualité que celle du LATTICE. C'est d'ailleurs cette qualité ainsi que son esprit d'innovation constant qui rendent en définitive optimiste quant à son évolution future, en dépit de l'incertitude relative que nous venons de relever et dont l'analyse SWOT elle-même se fait l'écho.



4 ● Analyse axe par axe

Axe 1 : Notions linguistiques et transferts entre langues

Nom du responsable : M. Benjamin FAGARD

Effectifs (en ETP)

Effectifs du thème en Équivalents Temps Plein	Au 30/06/2012	Au 01/01/2014
ETP d'enseignants-chercheurs titulaires	1,8	1,4
ETP de chercheurs des EPST ou EPIC titulaires	1,4	1,4
ETP d'autres personnels titulaires n'ayant pas d'obligation de recherche (IR, IE, PRAG, etc.)	0,5	0,5
ETP d'autres enseignants-chercheurs (PREM, ECC, etc.)	1,5	1,9
ETP de post-doctorants ayant passé au moins 12 mois dans l'unité	1	
ETP d'autres chercheurs des EPST ou EPIC (DREM, etc.) hors post-doctorants	1	1
ETP d'autres personnels contractuels n'ayant pas d'obligation de recherche	0	0
ETP de doctorants	5	
TOTAL	12,2	6,2

● Appréciations détaillées

Thème 1 : opération espace-temps

Les travaux sur l'espace analysent la variabilité typologique dans l'expression de l'espace et du temps. Ils se situent dans la continuation des travaux du LATTICE et s'appuient sur des études de corpus (annotation d'un corpus de 2 millions de mots). Le projet vise la création d'un corpus comparable à partir de supports multimedia dans la continuation des travaux menés par le LATTICE dans le cadre des projets Trajectoire (fédération TUL) et "De l'espace au temps" (Labex TransferS).

Il serait bon sur ce point de situer le projet par rapport aux travaux sur le temps et l'espace qui, en traitement automatique des langues, visent eux aussi l'annotation du temps et de l'espace (par exemple, schémas d'annotation eSpaceML, SpatialML, ISO(Space, ISO-TimeML).



Thème 2 : l'asymétrie Espace-Temps dans l'ancrage des événements

Ce thème prolonge les recherches de haut niveau, menées dans le cadre de l'opération « espace », autour du rapport entre l'espace et le temps, mais aussi, sur les plans syntaxique et pragmatique, sur les passerelles entre fonctions intra-prédicatives et périphériques. De l'opération 2.1. du précédent projet (« Unités prédicatives du discours »), ce thème récupère les travaux sur la structure argumentale du verbe, en particulier les études sur la distinction arguments/adjoints. Pour l'essentiel toutefois (les aspects « modélisation de construction dynamique du sens des unités prédicatives » et « modélisation informatique des constructions prédicatives »), les activités développées dans le cadre de l'opération « Unités prédicatives du discours » ont été intégrées à l'axe 3 du projet scientifique 2014-2018.

Centrer la recherche sur la distinction arguments/adjoints et sur les propriétés sémantiques des adjoints locatifs et temporels aboutit à la définition d'une problématique centrale et fédératrice dans les domaines syntaxique et sémantique. Pourquoi cependant ce thème n'intègre-t-il pas la dimension comparative (transfert d'une langue à l'autre) qui devrait unifier les thèmes au sein de l'axe 1 ? Il faudrait en effet envisager une approche comparative pour cette thématique. Elle s'y prête d'ailleurs très bien.

Thème 3 : Transferts de constructions d'une langue à l'autre

Le sous-thème « Transferts entre langues et transposition de textes » est une innovation par rapport aux travaux précédents du laboratoire. Il consistera en une comparaison des traductions en différentes langues européennes de textes littéraires français contemporains.

Eu égard à sa problématique ciblée sur des constructions particulières propres aux textes considérés, le projet est à la fois innovant et ambitieux. Le choix des langues cibles serait cependant à justifier par des considérations autres que strictement pratiques (spécialistes des langues choisies disponibles au sein de l'équipe) en raison de son impact probable sur les conclusions de l'enquête.

Le deuxième sous-thème « Transferts entre langues et constructions comparatives » prolonge l'opération « comparaison » qui est une des activités scientifiques marquantes du projet quadriennal 2010-2013 du LATTICE. L'ouverture de la recherche vers une approche comparative et typologique ne peut qu'accroître la validité des résultats obtenus et favoriser la reconnaissance internationale de l'équipe. Il faudrait toutefois à ne pas multiplier les axes de recherche (le projet envisage des développements sur le plan de la comparaison et de la typologie, de la diachronie, de la morphologie (structure du marqueur) et de la sémantique (rapport scalarité / comparaison).

Les points forts de cette recherche sont : (i) l'impact élevé de la recherche au sein du monde académique, (ii) la haute qualité des chercheurs impliqués, (iii) la mise en place d'une base de données spécialisée (essentiellement comme, aussi / autant...que).

Il reste que le développement du thème au sein du LATTICE est essentiellement lié aux travaux de chercheurs qui sont arrivés en fin de carrière académique. Il paraît donc raisonnable de réduire le nombre des orientations de recherche programmées dans le projet pour ce thème.

Thème 4 : Frame lexicale, culture et transferts

Le thème "Frame lexicale, culture et transferts" cible l'étude syntaxique et sémantique des verbes de don et d'échange dans différentes langues (français, anglais, nen et dalabon) visant en particulier à identifier la part culturelle des différences observées et leurs manifestations linguistiques. Outre l'étude contrastive, une étude diachronique est également envisagée au travers de collaborations internes au Lattice.



Conclusion

On constate dans ce premier axe un regroupement un peu hétéroclite des opérations 1.1. (« comparaison »), 1.2. (« espace ») et en partie 2.1. (« unités prédicatives du discours ») du contrat quadriennal en cours.

Les deux premières opérations ont produit des recherches de haute qualité, qui ont un impact scientifique important au sein du monde académique.

Parmi l'ensemble des thèmes de recherche du LATTICE, le dernier (« unités prédicatives du discours ») paraît par contre moins productif et fructueux. La qualité scientifique de la production 2009-2012 dans ce domaine a été plutôt moyenne.

La cohérence thématique et méthodologique de l'axe est à approfondir.



Axe 2 : Mécanismes de composition du discours

Nom du responsable : M. Frédéric LANDRAGIN

Effectifs (en ETP)

Effectifs du thème en Équivalents Temps Plein	Au 30/06/2012	Au 01/01/2014
ETP d'enseignants-chercheurs titulaires	3,9	3,3
ETP de chercheurs des EPST ou EPIC titulaires	2,2	2,2
ETP d'autres personnels titulaires n'ayant pas d'obligation de recherche (IR, IE, PRAG, etc.)	1	1
ETP d'autres enseignants-chercheurs (PREM, ECC, etc.)	0,5	1,1
ETP de post-doctorants ayant passé au moins 12 mois dans l'unité	1	
ETP d'autres chercheurs des EPST ou EPIC (DREM, etc.) hors post-doctorants	0	0
ETP d'autres personnels contractuels n'ayant pas d'obligation de recherche	0	0
ETP de doctorants	7	
TOTAL	15,6	7,6

• Appréciations détaillées

Cet axe se décline en quatre thèmes dont trois sont dans la continuité plus ou moins directe de recherches menées précédemment (zone préverbale, référence, marqueurs discursifs), tandis que le dernier (variations d'un genre à l'autre, contraste écrit/oral) correspond à une orientation nouvelle.

L'originalité de l'ensemble du projet se trouve dans la dimension contrastive qui est nettement déclarée pour mener à bien les futures études et analyses (différentes langues, différents genres, différents états du français, données orales/écrites).

Thèmes 1, 2 et 3 : Zone préverbale et structure informationnelle ; Référence ; Marqueurs discursifs

Ces trois thèmes se situent dans le prolongement des opérations 2.2 « Éléments initiaux, Ordre des mots, structure informationnelle et textuelle » et 2.3 « identification des référents et transitions référentielles », menées dans le plan de recherche précédent. Le second thème propose trois objets d'étude prometteurs : (i) une typologie des marqueurs de référence, (ii) l'étude de la construction de chaînes de coréférence, (iii) le rapport entre chaînes de coréférence et structures discursives tandis que les deux autres prolongent des recherches qui ont notablement contribué à la notoriété de l'unité dans le domaine des cadratifs, de l'ordre des mots et des connecteurs, le tout envisagé tant en synchronie qu'en diachronie.

Ces thèmes s'inscrivent également dans une des préoccupations majeures, et nouvelles, de l'axe 2, la variation selon le genre textuel, tout en prévoyant aussi une dimension comparative avec d'autres langues (cf. axe 1).



Étant donné d'autres besoins prioritaires, le LATTICE ne reprend plus dans ses objectifs prioritaires le recrutement d'un(e) psycholinguiste. Il vise plutôt à développer des coopérations avec d'autres unités de recherche. La décision semble cohérente au niveau des orientations globales du LATTICE mais laisse planer certaines incertitudes concernant l'apport psycholinguistique, crucial pour la validation des observations et des hypothèses linguistiques formulées dans le cadre des travaux sur la référence, les éléments initiaux ou les marqueurs discursifs.

On recommande donc de veiller à assurer rapidement des collaborations solides avec une équipe de psycholinguistes. Il faudrait aussi être attentif à gérer de manière cohérente la multiplicité des paramètres de variation (comparaison entre langues, dimension diachronique, la distinction entre l'oral et l'écrit, entre genres textuels) dans le développement de la recherche.

Thème 4 : Variations d'un genre à l'autre, contraste écrit/oral

Ce thème ouvre un champ de recherche jusqu'alors inexploré au LATTICE : la variation écrit/oral et les grammaires à mobiliser pour une description fine de leur interaction. Cette recherche sera menée sur le discours scientifique dans le cadre du projet EIIIDA (labex TransferS), il aura donc une dimension contrastive et portera essentiellement sur des faits de syntaxe.

Il serait intéressant de croiser cette recherche de nature purement synchronique avec les travaux sur la diachronie où le laboratoire s'est précédemment illustré. On a tout lieu de croire en effet que l'interaction entre l'oral et l'écrit est un facteur important de l'évolution des langues, qu'il serait donc utile de prendre aussi en compte.

Conclusion

- Avis global sur l'axe :

L'axe paraît particulièrement à même de mener des recherches significatives sur ces quatre thèmes grâce à la continuité de ses programmes et à leur expansion cohérente.

- Points forts :

Cet axe se positionne au cœur des apports scientifiques qui fondent la notoriété du LATTICE. Il exploite l'expertise accumulée en sémantique et en pragmatique et la connaissance des procédures méthodologiques de la linguistique de corpus. L'axe est en outre supporté par d'importants projets financés qui permettront également de renforcer les partenariats. On relève des implications dans plusieurs projets (entre autres, Labex TransferS) et l'obtention d'un projet ANR (ORFEO) dont un des membres de cet axe est porteur.

- Points à améliorer :

Si la dimension contrastive entre les langues semble bien être un dénominateur commun des travaux de l'axe 2, il est à regretter que celle-ci ne s'envisage qu'en fonction des forces de l'unité et des projets dans lesquels l'unité intervient (telles langues pour tel phénomène et telles autres langues pour tel autre phénomène), le choix des langues à comparer mériterait de s'appuyer sur une justification plus linguistique et l'analyse contrastive mériterait de devenir un vrai enjeu.

Il en est de même pour ce qui est de la notion de « genre », mentionnée de manière récurrente dans ces quatre opérations, qui mériterait d'être mieux problématisée, tant cette notion est loin d'être consensuelle dans la communauté des linguistes.

- Recommandations :

Il faudrait essayer d'articuler mieux les différents thèmes à l'intérieur de cet axe pour en expliciter la cohérence. Le questionnement théorique et critique des concepts pourrait favoriser une telle démarche d'éclaircissement.



Axe 3 : Corpus, modélisations et traitements automatiques

Nom du responsable : M^{ME} Isabelle TELLIER

Effectifs (en ETP)

Effectifs du thème en Équivalents Temps Plein	Au 30/06/2012	Au 01/01/2014
ETP d'enseignants-chercheurs titulaires	1,3	1,3
ETP de chercheurs des EPST ou EPIC titulaires	1,4	1,4
ETP d'autres personnels titulaires n'ayant pas d'obligation de recherche (IR, IE, PRAG, etc.)	0,5	0,5
ETP d'autres enseignants-chercheurs (PREM, ECC, etc.)	0	0
ETP de post-doctorants ayant passé au moins 12 mois dans l'unité	0	
ETP d'autres chercheurs des EPST ou EPIC (DREM, etc.) hors post-doctorants	0	0
ETP d'autres personnels contractuels n'ayant pas d'obligation de recherche	0	0
ETP de doctorants	2	
TOTAL	5,2	3,2

• Appréciations détaillées

Cet axe est une création du nouveau projet qui reprend, avec la nette volonté de les développer, des « opérations » du bilan précédent (TAL et linguistique outillée). Il se décline en trois thèmes (Recherches sur la construction dynamique du sens ; Modélisation de l'évolution des langues ; Traitements automatiques et applications). Les participants de cet axe sont également fortement impliqués dans le projet transversal "Humanités numériques" ainsi que l'opération transversale "Linguistique Outillée".

▪ Avis global sur l'axe :

Les recherches de l'axe "Corpus, modélisation et traitement automatiques" portent globalement sur la modélisation de l'évolution des langues, l'extraction de connaissances à partir de corpus, l'extraction d'information, l'apprentissage supervisé et la construction du sens dans une perspective de philosophie du langage. La création d'un nouvel axe visant le développement de méthodes de modélisation informatique des données langagières est légitimée à la fois par l'arrivée récente dans le laboratoire de deux chercheurs de haut niveau en TAL, aux profils très complémentaires, et par l'intérêt manifesté par d'autres thématiques du laboratoire en faveur du développement de méthodes informatiques permettant à la fois d'aider l'annotation et de l'exploiter par des méthodes d'apprentissage. Les compétences des chercheurs en TAL dans le domaine de l'extraction d'informations et de l'apprentissage automatique constituent de fait un atout majeur pour le laboratoire.



- Points forts et possibilités liées au contexte :

L'émergence d'activités en apprentissage automatique complètera de façon pertinente l'expérience développée dans l'équipe en matière de corpus annotés. L'implication du laboratoire dans le master PluriTAL, composante de Paris 3, constitue un atout important pour le développement de l'axe. Ce point devrait être plus nettement mis en valeur dans le projet, en spécifiant plus précisément la nature des interactions entre formation et recherche. Le projet laisse présager un renforcement positif des contacts avec les entreprises dans le champ du TAL.

- Points à améliorer et risques liés au contexte :

Le projet de cet axe laisse craindre néanmoins un certain éparpillement des forces en TAL. Il est composé de sous-thématiques hétérogènes dont les points de convergence sont pour l'instant encore à définir : aux côtés d'une thématique clairement définie en lien avec les spécialités des chercheurs en TAL et bien articulée avec les thématiques du laboratoire (traitements de corpus : annotation, extraction, acquisition), les deux thématiques "construction dynamique du sens" et "modélisation de l'évolution des langues" semblent encore embryonnaires.

Des recrutements en TAL (chercheurs/ingénieurs) seront certainement nécessaires pour garantir la faisabilité du projet. Des besoins en termes d'infrastructure et d'équipement informatique sont à anticiper pour garantir la réalisation de projets impliquant des traitements informatiques complexes ; ces besoins sont en partie compensés par une très bonne stratégie de partenariats.

- Recommandations :

Vu les effectifs limités du LATTICE et le nombre restreint d'informaticiens et de linguistes, il semblerait judicieux soit de resserrer les objectifs scientifiques (par exemple, en restreignant l'apprentissage supervisé à une ou deux applications spécifiques) soit de renforcer les recrutements en particulier par le biais de projets.

Les thématiques annoncées doivent être resserrées de manière à rechercher des convergences au sein de l'axe. Un effort de formation des chercheurs de l'équipe aux méthodes d'apprentissage automatique est souhaitable, pour les rendre aussi autonomes que possible sur ces activités.



5 • Déroulement de la visite

Date de la visite :

Début : le jeudi 6 décembre 2012 à 9h30

Fin : le jeudi 6 décembre 2012 à 18h15

Lieu de la visite :

Institution : UMR 8094 LATTICE (CNRS-ENS-U. Sorbonne Nouvelle)

Adresse : 1, rue Maurice Arnoux 92120 Montrouge

Locaux spécifiques visités :

Locaux de l'unité de recherche (bureaux, salles de réunion, secrétariat)

Déroulement de la visite :

9h30-10h30 : réunion à huis-clos du comité d'experts

10h30-11h45 : Début de la réunion plénière

- présentation du comité et de la démarche de l'évaluation par la présidente
- tour de table des présents pour la présentation : direction de l'unité de recherche, membres de l'UR, personnel administratif, doctorants
- Présentation du bilan de l'UR par la direction
- Présentation du bilan et du projet des axes par leur responsable

Déjeuner 12h10-13h45

Reprise visite : 13h45

- 13h45-15h55 : Suite de la réunion plénière
- Rencontre du comité d'experts avec les représentants des tutelles
- Présentation du projet de l'UR par la direction
- Présentation du projet des axes par leur responsable

Huis-clos du comité

- Rencontre du comité d'experts avec les doctorants seuls
- Rencontre du comité avec le personnel
- Rencontre du comité avec la direction de l'unité de recherche

Fin des entretiens : 16h45

16h45-17h45 : Réunion à huis clos du comité d'experts.



Points particuliers à mentionner

La visite s'est déroulée selon le programme prévu. Seule la dernière rencontre avec la direction de l'UR n'était pas planifiée, mais elle a pu avoir lieu, grâce à la disponibilité du directeur et de la directrice-adjointe, à la demande du comité d'experts qui a souhaité clarifier certains points évoqués rapidement pendant les réunions plénières. Ce dernier entretien a duré une vingtaine de minutes.



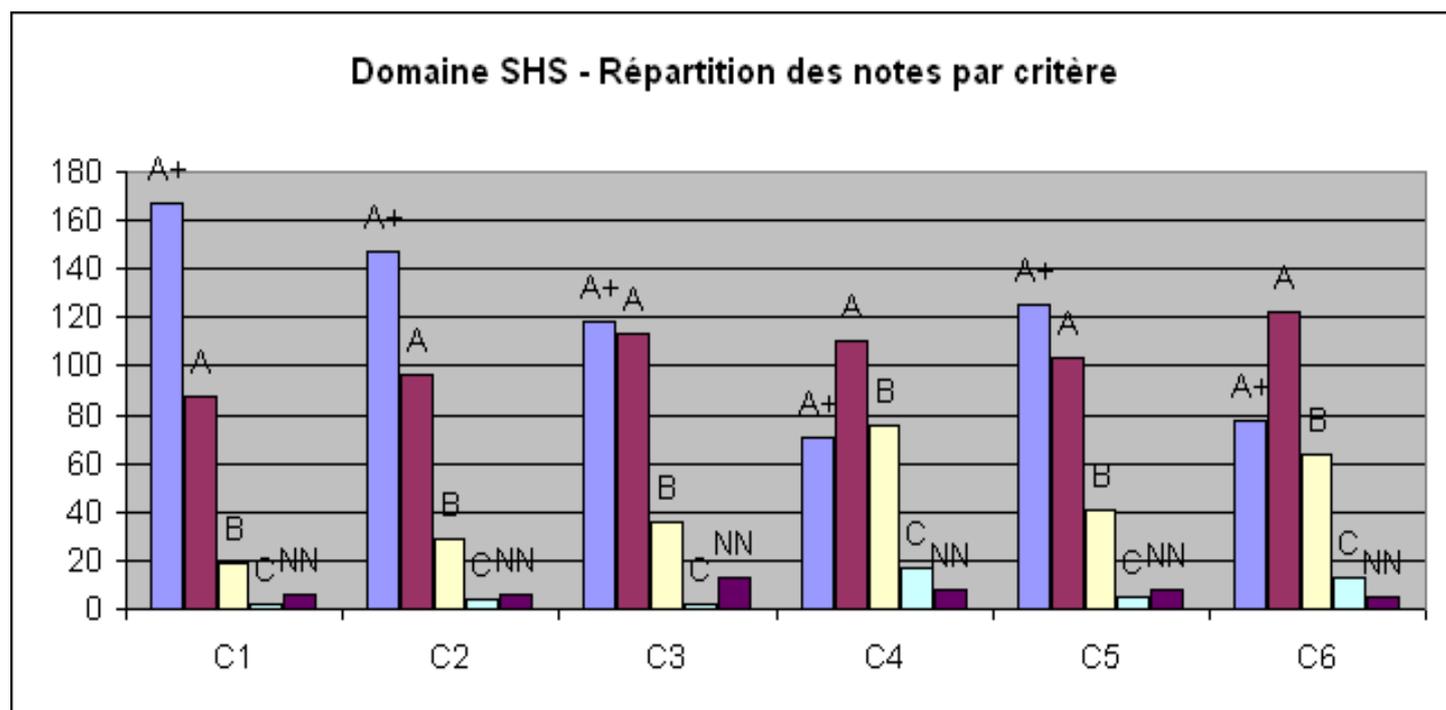
6 • Statistiques par domaine : SHS au 10/06/2013

Notes

Critères	C1 Qualité scientifique et production	C2 Rayonnement et attractivité académiques	C3 Relations avec l'environnement social, économique et culturel	C4 Organisation et vie de l'entité	C5 Implication dans la formation par la recherche	C6 Stratégie et projet à cinq ans
A+	167	147	118	71	125	78
A	88	96	113	110	103	122
B	19	29	36	76	41	64
C	2	4	2	17	5	13
Non Noté	6	6	13	8	8	5

Pourcentages

Critères	C1 Qualité scientifique et production	C2 Rayonnement et attractivité académiques	C3 Relations avec l'environnement social, économique et culturel	C4 Organisation et vie de l'entité	C5 Implication dans la formation par la recherche	C6 Stratégie et projet à cinq ans
A+	59%	52%	42%	25%	44%	28%
A	31%	34%	40%	39%	37%	43%
B	7%	10%	13%	27%	15%	23%
C	1%	1%	1%	6%	2%	5%
Non Noté	2%	2%	5%	3%	3%	2%





7 • Observations générales des tutelles

Paris, le 15 avril 2013

Thierry Poibeau
Directeur de recherche au CNRS
Directeur du laboratoire LATTICE (UMR 8094)

The logo for the Lattice laboratory, featuring the word "Lattice" in a stylized font with a grid pattern behind it.

Langues, Textes,
Traitements Informatiques, Cognition

www.lattice.cnrs.fr

ENS
1, rue Maurice Arnoux
92120 Montrouge

T. 01 58 07 66 20
F. 01 58 07 66 29
prenom.nom@ens.fr
UMR 8094

Objet : Remarques du laboratoire Lattice (UMR 8094) suite à la visite du comité d'experts de l'AERES du 6 décembre 2012

Les membres du laboratoire Lattice ont pris connaissance du projet de rapport faisant suite à la visite du comité d'experts de l'AERES qui a eu lieu le 6 décembre dernier.

Nous remercions les membres du comité pour leur écoute, leurs questions et surtout leurs remarques dont nous tiendrons le plus grand compte. Nous sommes très heureux de voir reconnues la visibilité du laboratoire, la qualité de la production scientifique, l'importance des partenariats et la capacité de renouvellement méthodologique.

D'autres remarques concernent des risques potentiels dont nous avons conscience. Nous souhaitons simplement apporter ici quelques précisions sur le contenu du rapport.

Le comité a souligné la nécessité d'assurer une « cohérence thématique des différents travaux menés dans l'unité ». Nous veillerons bien évidemment à ce point qui est effectivement crucial. Cette cohérence nous semble cependant déjà relativement présente : le laboratoire s'intéresse à des phénomènes linguistiques précis (essentiellement liés à la syntaxe, à la sémantique et au discours) avec une approche méthodologique claire (tous les projets s'appuient sur la linguistique outillée et/ou la psycholinguistique). Nous envisageons maintenant d'élargir progressivement le nombre de langues abordées – en veillant à leur diversité en fonction des phénomènes étudiés – et d'approfondir les recherches en traitement automatique du langage naturel. La plupart des opérations de recherche mêlent des membres du laboratoire avec des compétences complémentaires, contribuant ainsi à la cohésion d'ensemble.

Le comité indique également que certaines nouvelles orientations de recherche « pèchent parfois par une certaine abstraction ou une trop grande généralité ». Il faut rappeler que le laboratoire s'est considérablement renouvelé durant le précédent quadriennal, ce qui a permis de maintenir l'effectif et le potentiel de recherche intacts. Cet important renouvellement a amené à proposer un projet de recherche largement remanié, avec des opérations de recherche qui se poursuivent mais aussi d'autres qui portent sur des thèmes nouveaux. Certaines opérations marginales dans le projet précédent ont été intégrées explicitement, afin de mieux rendre compte de l'ensemble des recherches menées au laboratoire. Enfin, le Lattice assume une certaine prise de risque liée au lancement de nouvelles opérations, dans la mesure où il nous semble

Sous la tutelle du



important d'explorer des voies nouvelles et originales. Les brefs descriptifs fournis (du fait des contraintes de longueur imposées) peuvent sans doute apparaître un peu abstraits mais ces opérations de recherche s'appuieront sur des données attestées et conduiront donc très rapidement à des recherches très concrètes. Un bilan interne sera fait pour chaque opération à mi-parcours du prochain contrat quinquennal et les opérations qui n'auront pas abouti à des résultats tangibles seront arrêtées le cas échéant.

Les interrogations concernant la place du Lattice vis-à-vis de ses tutelles ne nous semblent pas totalement fondées. Malgré les départs en retraite, le nombre d'enseignants-chercheurs liés à l'Université Paris 3 Sorbonne Nouvelle est resté stable. Notons que J.-M. Debaisieux fait partie du bureau de l'Ecole doctorale 268 et qu'I. Tellier est fortement impliquée dans la formation Plurital (master de traitement automatique des langues commun à Paris 3, à Paris 10 et à l'INaLCO). La quasi totalité des doctorants du Lattice font partie de l'ED268 et le laboratoire accueille chaque année plusieurs stagiaires de Master de l'Université Paris 3. Pour le reste, les liens se sont aussi renforcés avec l'Ecole normale supérieure (participation au labex Transfers, à l'ED540 et à l'offre de cours de l'Ecole) ainsi qu'avec le CNRS (I. Tellier est chargée de missions à l'INS2I, A.-M. Argenti siège au sein de la section 34 pour la mandature 2013-2016). Nous continuerons bien évidemment à veiller à avoir un lien fort avec nos différentes tutelles.

Le comité d'experts s'interroge également sur la stratégie concernant la psycholinguistique. La direction du laboratoire n'a pas fait passer cette activité au second plan mais il faut bien constater la difficulté à recruter sur ce thème malgré des demandes récurrentes. Il nous a donc semblé que des partenariats ciblés étaient la meilleure solution pour aller de l'avant, plutôt que d'attendre passivement un éventuel recrutement. Le laboratoire continuera de soutenir l'approche psycholinguistique (dont nous mesurons l'importance vu la nature de notre projet de recherche) en affichant ce thème comme axe prioritaire, notamment au niveau des recrutements.

Enfin, le rapport souligne à juste titre le cas d'un personnel ITA spécialement dévoué, dont le statut est très inférieur aux fonctions occupées. La direction du laboratoire en a conscience et, d'une manière générale, essaie d'assurer au mieux une évolution de carrière satisfaisante pour les personnels ITA grâce notamment à la formation continue. Mais il faut rappeler que la direction du laboratoire n'a en définitive que peu de prise sur l'avancement des personnels. Nous avons demandé un avancement accéléré pour la personne en question les années passées et nous le ferons jusqu'à obtenir satisfaction (un changement de grade a été obtenu en 2012 mais nous avons conscience qu'un changement de corps correspondrait mieux aux tâches effectuées). Gageons que le rapport AERES sera une aide précieuse de ce point de vue.

En conclusion, nous souhaitons à nouveau remercier l'ensemble du comité pour cette évaluation très détaillée, qui nous permet d'avoir un regard extérieur précieux pour améliorer la qualité de notre recherche.

Thierry Poibeau

