

Appel à projets franco-allemand
en sciences humaines et sociales

Ausschreibung eines deutsch-
französischen Programms in den Geistes-
und Sozialwissenschaften

Programme non-thématique

Ohne thematische Vorgaben

2012

2012

Titre du projet
**L'évolution du système prépositionnel
du français**

Projekttitel
**Die Entwicklung des französischen
Präpositionalsystems**

Acronyme

Kennwort

PRESTO

PRESTO

Noms des deux coordinateurs
(français et allemand) du projet et
institutions de rattachement

Peter Blumenthal, Romanisches Seminar,
Universität zu Köln
Denis Vigier, ICAR (UMR 5191),
Université Lumière, Lyon 2

Namen der beiden Projektleiter
(auf deutscher und auf französischer
Seite) und ihrer Institutionen

INHALTSVERZEICHNIS

I. DARSTELLUNG DES PROJEKTES

1. ALLGEMEINE ANGABEN	
1.1. Antragsteller, Projektbeteiligte und Kooperationspartner	3
1.2. Thema des Vorhabens und Akronym	5
1.3. Fach- und Arbeitsrichtung	5
1.4. Antragszeitraum	5
1.5. Zusammenfassung,.....	5
2. STAND DER FORSCHUNG, EIGENE VORARBEITEN	
2.1. Stand der Forschung	
2.1.1. Das Präpositionalsystem des Gegenwartsfranzösischen	6
2.1.2. Das Präpositionalsystem des Französischen in der Diachronie : Bestandsaufnahme	8
2.1.3. Der Begriff der nicht-freien Wortverbindung (<i>figement</i>)	9
2.2. Eigene Vorarbeiten zum Thema	
2.2.1. Arbeiten der deutschen Forschergruppe	10
2.2.2. Arbeiten der französischen Forschergruppen	12
3. ZIELE UND ARBEITSPROGRAMM	
3.1. Ziele	
3.1.1. Wissenschaftliche Ziele	13
3.1.2. Zu erwartende Synergien	18
3.1.3. Veröffentlichung der Ergebnisse	19
3.2. Arbeitsprogramm, Methoden, Zeitplan	
3.2.1. Arbeitsprogramm und Methoden	
3.2.1.1. Programm und Methodik der sprachwissenschaftlichen Untersuchung	21
3.2.1.2. Programm und Methoden hinsichtlich der Korpuserstellung und –auswertung	22
3.2.2. Zeitplan und Aufteilung der Arbeitspakete nach Forschergruppen	
3.2.2.1. Programm und Methodik der sprachwissenschaftlichen Untersuchung	26
3.2.2.2. Programm und Methoden hinsichtlich der Korpuserstellung und –auswertung	27
3.3. Literaturverzeichnis zum Stand der Forschung, zu den Zielen und dem Arbeitsprogramm	30
4. BEANTRAGTE MITTEL /MODULE	36
5. ÜBERBLICK ÜBER DIE BEANTRAGTEN MITTEL	38
II. ANHANG	
Lebensläufe aller Beteiligten	40
Anhänge zur Veranschaulichung der Werkzeuge, Methoden und Zielsetzungen des Projekts	

1. Allgemeine Angaben

1.1. Antragsteller, Projektbeteiligte und Kooperationspartner

1.1.1. Namen, Adressen und Umfang des zeitlichen Engagements der Antragsteller auf deutscher und französischer Seite

Name	Vorname	Adresse	E-mail	Umfang des zeitlichen Engagements
BLUMENTHAL	Peter	Universität zu Köln	peter.blumenthal@uni-koeln.de	50%
VIGIER	Denis	Université Lumière Lyon2 & UMR ICAR 5191	Denis.Vigier@univ-lyon2.fr	50%

1.1.2. Wissenschaftlerinnen und Wissenschaftler, mit denen für dieses Vorhaben eine konkrete Vereinbarung zur Zusammenarbeit auf deutscher und auf französischer Seite besteht

a. Deutschland

Name	Vorname	Adresse	E-mail	Umfang des zeitlichen Engagements
DIWERSY	Sascha	Universität zu Köln	sascha.diwersy@uni-koeln.de	30%
GILLOV	Patrick	Universität zu Köln	pgililov@uni-koeln.de	50%

b. Frankreich

Name	Vorname	Adresse	E-mail	Umfang des zeitlichen Engagements
CHAROLLES	Michel	Paris 3 & UMR LaTTICe	michel.charolles@ens.fr	10 %
DEMONET	Marie-Luce	Université F. Rabelais, Tours & UMR 6576 (CESR)	marie-luce.demonet@univ-tours.fr	10%
FAGARD	Benjamin	CNRS & UMR LaTTICe	benjamin.fagard@ens.fr	20 %
GARDELLE	Laure	ENS & UMR ICAR 5191	laure.gardelle@ens-lyon.fr	10%
GUILLOT	Céline	ENS & UMR ICAR 5191	celine.guillot@ens-lyon.fr	20%
HEIDEN	Serge	ENS & UMR ICAR 5191	slh@ens-lyon.fr	20%
LAVRENTEV	Alexei	CNRS & UMR ICAR 5191	alexei.lavrentev@ens-lyon.fr	20%
LEEMAN	Danielle	CNRS & UMR MODYCO 7114	danielle.leeman@univ-paris10.fr	30%
PINCEMIN	Bénédicte	CNRS & UMR ICAR 5191	benedicte.pincemin@ens-lyon.fr	20%
SARDA	Laure	CNRS & UMR LaTTICe	laure.sarda@ens.fr	10 %

1.1.3. Kooperationspartner (außerhalb des beantragten Projekts)

Für die Analyse von Präpositionen:

Dany Amiot, *Professeur des Universités, Université de Lille 3*

Jacques François, *Professeur des Universités émérite, Université de Caen*

Jürgen Lang, *Professeur des Universités émérite, Université d'Erlangen*

Walter de Mulder, *Professeur des Universités, Université d'Anvers*

Christoph Schwarze, *Professeur des Universités émérite, Université de Constance*

Achim Stein, *Professeur des Universités, Université de Stuttgart, Institut für Linguistik Romanistik.*

Dejan Stosic, *Maître de Conférences, Université d'Artois*

Für den diachronen Ansatz:

Bernard Combettes, *Professeur des universités émérite, Université Nancy2*

Christiane Marchello-Nizia, *Professeur des Universités émérite, ENS Lyon*

Nathalie Fournier, *Professeur des Universités, Université Lyon 2*

Sophie Prévost, *Chargée de Recherche CNRS, UMR LaTTICe*

Sylviane Rémi-Giraud, *Professeur Emérite des Universités, Université Lyon2.*

Für die nicht-freie Wortverbindung (figement):

Gaston Gross, *Professeur des Universités émérite, Université de Villetaneuse (Paris 13)*

Béatrice Lamiroy, *Professeur des Universités, Université de KU Leuven*

Für Lexikometrie und Textometrie, automatische Sprachverarbeitung

Dominique Labbé, *Professeur des Universités émérite, PACTE (UMR CNRS/IÉP/UPMF/UJF 5194) Grenoble*

Benoît Sagot, *chercheur INRIA, au sein du projet Alpage (INRIA Paris Rocquencourt et Université Paris 7), France. <http://alpage.inria.fr/~sagot>*

Jean-marie Viprey, *Professeur des Universités, Université de Franche- Comté*

Für die Korpuslinguistik:

Martin Holmes, *Ingénieur de recherche, University of Victoria Humanities Computing and Media Centre, Canada*

F. Neveu, *Professeur des Universités, Université de Caen*

R. Morrissey, *PhD in Romance Languages and Literatures, University of Chicago (directeur de l'ARTFL) <http://rll.uchicago.edu/faculty/morrissey>*

M. Olsen *ARTFL / University of Chicago (directeur adjoint de l'ARTFL)*

1.2. Thema des Vorhabens und Akronym

Thema: Die Entwicklung des französischen Präpositionalsystems

Akronym: PRESTO

1.3. Fach- und Arbeitsrichtung

Französische Sprachwissenschaft (Korpuslinguistik, Textometrie, Syntax, Sprachgeschichte, Theorie des Sprachwandels)

1.4. Antragszeitraum

36 Monate

1.5. Zusammenfassung

Ziel des Projektes ist die diachrone Erforschung der Verwendungen und der semantischen sowie diskursiven Werte der französischen Präpositionen *à, en, par,*

contre, dès, devant, entre, pour, sans, sur, sous, vers, dans von dem ältesten Sprachzustand¹ (9. Jh.) bis zum Französischen der Gegenwart.

Gegenstand der von einem statistisch-distributionellen Ansatz ausgehenden, software-gestützten Untersuchung sind Veränderungen des kombinatorischen Verhaltens von Präpositionen in Abhängigkeit von Epoche (diachrone Entwicklung), Textgattung und Autor. Unser Augenmerk wird sich dabei besonders auf die Erfassung hochspezifischer Kombinationen² mit den genannten Präpositionen richten. So werden wir Emergenz, Umgestaltung und Schwinden mehr oder weniger fester (nicht-freier) Präpositionalsyntagmen untersuchen. Wir hoffen, auf der Grundlage eines neuen Analysemodells zu einem neuartigen Verständnis der historischen Entwicklung in diesem Bereich zu gelangen; dieses Modell verbindet zwei methodische Verfahrensweisen, nämlich (1) die Untersuchung des Zusammenhangs zwischen Quantitativem (Verwendungen) und Qualitativem (Werte) sowie (2) die Untersuchung der Prinzipien, welche diachron das Streben der Sprecher nach semantischen und diskursiven Mehrwerten als regelrechten Motor sprachlichen Wandels gesteuert haben.

Eine der Voraussetzungen – zugleich aber auch einer der wichtigsten Erträge - des Projektes wird in der Zusammenstellung eines lemmatisierten und annotierten Korpus bestehen, in dem alle Perioden der französischen Sprachgeschichte sowie verschiedene Diskursarten und Texttypen vertreten sind. Dieses gemeinsam von der französischen und der deutschen Forschergruppe zu erstellende Korpus kann sich teilweise auf die schon verfügbaren Textressourcen aus Lyon (Mittelalter), Tours (16. Jh.), Nancy (ATILF) und Köln (Moderne, 17. – 20. Jh.) stützen. Gemeinsames Ziel beider Gruppen ist es, die Öffnung eines größtmöglichen Teils des Korpus zu ermöglichen.

2. Stand der Forschung, eigene Vorarbeiten

2.1. Stand der Forschung

2.1.1. Das Präpositionalsystem des Gegenwartsfranzösischen

Die Arbeiten der letzten zwanzig Jahre über Präpositionen (vgl. die Bibliografien von Vaguer 2006, 2007, de Mulder & Stosic, 2009), haben dazu beigetragen, die Herangehensweise an diese Wortart auf syntaktischer, semantischer und distributioneller Ebene zu erneuern.

2.1.1.1. Auf syntaktischer Ebene haben zahlreiche neuere Arbeiten (u. a. Melis, 2003; Leeman, 2006, 2008) gezeigt, dass die traditionell vorgebrachten Kriterien zur Definition der Präposition (Vorhandensein eines einzigen, nachgestellten, obligatorischen Rektum nominaler Natur, auf das die Präposition Zwänge semantischer, syntaktischer und distributioneller Art ausüben kann) nicht immer zutreffen. An dieser Stelle sei nur auf einen Punkt hingewiesen, der in unmittelbarem Zusammenhang mit dem Projekt steht: Le Pesant 2006, Leeman 2006, 2008, u. a. haben gezeigt, dass sich Präpositionen syntaktisch über ihre Fähigkeit zur Subkategorisierung ihrer Argumente (NP, Satz, Infinitivphrase, Adverbialphrase,

¹ Mit Ausnahme von *dans* sind all diese Präpositionen seit dem 9. oder 11. Jh. in Gebrauch (vgl. Marchello-Nizia, 1999: 119-120). Die Präposition *dans*, für die sich erst später (ab dem 12. Jh.) Belege finden lassen, ist wegen ihrer Konkurrenz zu *en* (Gegenstand einer gesonderten Untersuchung innerhalb des Projektes, vgl. § 3.2.1.1) ins Inventar aufgenommen worden. Die Präposition *de*, deren automatische Erkennung erhebliche Schwierigkeiten aufwirft, wurde ausgeschlossen.

² Es handelt sich hierbei um Kookkurrenzen, deren relativ hohe Frequenz nicht dem Zufall geschuldet sein kann und die in wahrscheinlichkeitstheoretischen Berechnungen als „spezifisch“ ausgewiesen wird.

Adjektivalphrase, Präpositionalphrase oder Nullkomplement) definieren lassen. Da nicht alle dieselben kombinatorischen Spielräume aufweisen, lässt sich eine Einteilung gemäß dem Kriterium der Auswahlbeschränkungen vornehmen, welche sie im Prinzip³ den Kategorien ihrer Rekta auferlegen; diese Einteilung kann in Form von syntaktischen Übersichtstafeln dargestellt werden (Vaguer, ebenda). Eines der Ziele des Projektes wird darin bestehen, solche Übersichtstafeln für vordefinierte Zeitschnitte, die den Zeitraum vom 16. bis zum 20. Jh. abdecken, anzufertigen.

2.1.1.2. a) Auf semantischem Gebiet haben nach der kognitiven Wende, welche die Präpositionen-Forschung während der 80er Jahre genommen hat (Vandeloise 1986, Lakoff 1987, Langacker 1987, usw.), Ansätze einer instruktionellen (Anscombe 1991, Berthonneau & Cadiot 1993, Cadiot 1997, usw.), konstruktivistischen⁴ (Cadiot & Visetti 2001, Franckel 2002, Franckel & Paillard 2008) oder auch konstruktionellen (Amiot & de Mulder 2012) Linguistik neue Perspektiven eröffnet und neue Debatten angestoßen. Doch weisen diese neueren Ansätze im Allgemeinen einen erheblichen Nachteil auf: Die Formulierungen, zu denen sie führen (Instruktionen, Motive, schematische Formen), erweisen sich häufig als wenig überprüfbar und/oder zu „erklärungsmächtig“ (G. Kleiber, 1999: 47-49).

b) Die Arbeiten von D. Leeman (1995, 1998, 2008), die im Gegensatz zum bisweilen recht spekulativen Charakter von manchen jener Untersuchungen stehen, haben die Notwendigkeit aufgezeigt, beim Versuch, die semantische Identität einer Präposition zu bestimmen (ganz gleich auf welches theoretische Modell man sich dabei stützt), die distributionellen Zwänge zu berücksichtigen, welche die Präposition insbesondere ihrem Rektum auferlegt. So zeigt Leeman beispielsweise für die Präposition *en*, wie die Untersuchung der möglichen und unmöglichen Verbindungen der Präposition mit nominalen Argumenten es ermöglicht, schrittweise eine Hypothese über eine semantische Invariante aufzubauen.

c) Allerdings sagt die Berücksichtigung verbotener vs. erlaubter Kombinationen bei einer Präposition nichts über die Rangfolge der kombinatorischen Präferenzen aus, welche sich in den Gebrauchsarten beobachten lassen und welche einzig das Heranziehen umfangreicher Korpora zu Tage fördern kann. Ebendiese kombinatorischen Präferenzen (ob sie nun „frei“ oder gebunden sind) ermöglichen es aber, die tiefer liegenden⁵ Affinitäten zwischen der Präposition und ihren Kollokatoren zu erforschen.

2.1.2. Das Präpositionalsystem des Französischen in der Diachronie: Bestandsaufnahme

2.1.2.1 Bei den seit dem 19. Jh. erfolgten Untersuchungen zu französischen Präpositionen lassen sich zwei Hauptausrichtungen unterscheiden: zum einen synchrone Untersuchungen, die nur indirekt (d.h. durch Vergleich) Auskunft über die Entwicklung geben, zum anderen die wirklich diachronen Untersuchungen, die grundsätzlich nach der Entwicklung der analysierten Formen fragen. Seit etwa 20 Jahren finden sich zudem zahlreiche Untersuchungen, die im Rahmen der Grammatikalisierungstheorie erfolgen und in Anlehnung an die Arbeiten von Kilroe (1989) zugleich einen morphosyntaktischen und einen semantischen Ansatz verfolgen.

2.1.2.2. Synchrone Untersuchungen zu älteren Sprachzuständen

³ Wohlgermerkt *im Prinzip*, denn außer der Präposition kann auch der „externe Kopf“ (Melis, ebenda) der Präpositionalphrase (Verb, Adjektiv) auf die auf dem Rektum lastende Auswahlbeschränkung Einfluss nehmen.

⁴ S. hierzu E. Spang-Hanssen (1993: 20-24), C. Vaguer (2006: 119-122).

⁵ „Tiefer liegend“, da dem Bewusstsein der Sprecher nicht zugänglich.

Am zahlreichsten sind synchrone Untersuchungen. Sie beziehen sich auf verschiedene Perioden des Französischen – Alt- und Mittelfranzösisch, vorklassisches und klassisches Französisch – sowie auf unterschiedliche – semantische, syntaktische und morphosyntaktische - Charakteristika von Präpositionen.

Zunächst ist auf das Bestehen sehr feingradiger, im Allgemeinen auf der Grundlage eingeschränkter Korpora durchgeführter Analysen zum Gebrauch von Präpositionen bei dem ein oder anderen älteren Autor hinzuweisen. Die meisten dieser Untersuchungen stammen vom Ende des 19. Jhs. oder Anfang des 20. Jhs. (vgl. Michael, 1904; Shears, 1922; Bekkers, 1931 ...). Trotz ihres Alters stellen sie aufgrund ihrer philologischen Genauigkeit einen Ausgangspunkt von unschätzbarem Wert für moderne Untersuchungen dar.

Ebenfalls finden sich Untersuchungen zur morphosyntaktischen Funktionsweise von Präpositionen, insbesondere im Rahmen des generativistischen Ansatzes, wie Lebel (2003). Diese Untersuchungen haben es beispielsweise ermöglicht, Phänomene der Überdeckung zwischen Partikeln, Verbpräfixen und Präpositionen herauszustellen, die des Weiteren in Grammatiken beschrieben werden: Buridant (2000) für das Altfranzösische, Marchello-Nizia (1979/1997) für das Mittelfranzösische, Gougenheim (1974) für das klassische Französisch. Diese Grammatiken liefern im Allgemeinen eine relativ detaillierte Beschreibung einer kleinen Auswahl von Präpositionen, ohne sich an weniger frequenten oder weniger grammatikalisierten Präpositionen, noch an zusammengesetzten Präpositionen aufzuhalten.

2.1.2.3. Diachrone Untersuchungen

Einige Autoren haben sich mit der Entwicklung des Präpositionalsystems im eigentlichen Sinne befasst. So hat Marchello-Nizia (2002) die zunehmende Spezialisierung von Morphemen gezeigt, die im Altfranzösischen sowohl als Präpositionen als auch als Partikeln gebraucht werden und die dann allmählich ihre Fähigkeit verlieren, als Partikeln funktionieren zu können. Der morphosyntaktische Status von Präpositionen ist auch von Amiot & De Mulder (2002) untersucht worden, welche im Zusammenhang mit der Grammatikalisierungstheorie nach der Direktionalität der Entwicklung Adverb/Präfix/Präposition fragen. Der Frage nach der Grammatikalisierung von Präpositionen geht ganz allgemein auch Fagard (2006) nach, der die bei französischen Präpositionen beobachteten Grammatikalisierungstypen näher bestimmt. Die spezifischere Frage nach der Art und Weise, wie zusammengesetzte Präpositionen geschaffen werden, ist vornehmlich in Fagard & De Mulder (2007) behandelt worden. Shyldkrot & Kemmer (1995) untersuchen ihrerseits die Entwicklung der Gebrauchsarten der funktionalen Präpositionen *à* und *de* in Zusammenhang mit Transitivity. Unter einem ganz anderen Gesichtspunkt wird die Transitivity von Präpositionen von Dufresne, Dupuis & Tremblay (2003) untersucht. Zahlreiche Präpositionen sind in ihrer semantischen Entwicklung genau beschrieben worden: *à* (Kilroe 1989, 1994, Martineau 2000), *contre* (Dendale & De Mulder 1998, De Mulder & Vanderheyden 2001), *sur* (De Mulder & Vanderheyden 2002), *par* (Raithel 1875), *après* (Fagard 2004), *puis, depuis* (Fagard 2007), *devant* (Fagard & De Mulder 2010), *côté* (Fagard 2008), *dans* (Fagard & Sarda 2009).

Jedoch wurden bislang keine statistischen Untersuchungen an umfangreichen Korpora durchgeführt, um die diachronen Veränderungen der präferentiellen Kombinatorik von Präpositionen zu erfassen. Derartige Daten würden es ermöglichen, die diachrone Entwicklung der Phraseologie zu untersuchen, ein Thema, zu dem aktuell sehr wenige Studien vorliegen. Sie würden ebenso den Weg bereiten für ein innovatives Nachdenken über die „semantischen Mehrwerte“, die am Anfang jener Entwicklungen standen (s. u. 3.1.1.).

2.1.3. Der Begriff der nicht-freien Wortverbindung (*figement*)

Der komplexe und folglich schwierig zu definierende Begriff der nicht-freien Wortverbindung (auch „Fixierung“) hat in den letzten Jahren Anlass zu zahlreichen Veröffentlichungen in der französisch- und englischsprachigen Fachliteratur geboten (u. a. G. Gross 1996, Svensson 2004, Lecolle & Leroy 2006, Bolly 2008, Blumenthal & Mejri 2008, Lamiroy & al. 2010; Everaert 1995, Hudson 1998, Moon 1998, Hunston & Francis 2000, Wray 2002). Bei aller Vielfalt von Ansätzen und Konzeptionen herrscht Einigkeit darüber, dass es sich um ein zentrales und vielschichtiges Phänomen der Sprache handelt - sowohl auf qualitativer als auch auf quantitativer Ebene - (vgl. Dannell 1992, Senellart 1998). Vielfach wird seine psycholinguistische (Gedächtniseintragung: Grunig, 1997) und soziale (wichtige Rolle der Institutionalisierung: Moon, 1998) Dimension betont.

Unter den Kriterien, die angeführt werden, um zu bestimmen, ob eine Sequenz fixiert ist oder nicht (und wenn ja, bis zu welchem Grad), erfreuen sich derzeit drei eines breiten Konsenses: i) das Kriterium der semantischen Nicht-Kompositionalität, ii) das lexikalische Kriterium der paradigmatischen Uersetzbarkeit, iii) das Kriterium der Nicht-Modifizierbarkeit der morphosyntaktischen Markierungen. Zu diesen drei Kriterien kommt noch der grundsätzlich skalare Charakter der Fixierung, woraus sich die wohlbekannte Schwierigkeit ergibt, angesichts nicht-prototypischer Sequenzen (d.h. Sequenzen, die nicht die Gesamtheit der genannten Kriterien in sich vereinen) zu bestimmen, ob man es mit einem festen Ausdruck zu tun hat oder nicht. Im Übrigen wäre es sinnvoll von doppelter Skalarität zu sprechen, denn sie trifft gleichermaßen auf die Synchronie wie auf die Diachronie zu. In der Tat sind die (weder notwendigen noch hinreichenden (Klein & Lamiroy 2005)) Merkmale der Fixierung je nach betrachtetem Ausdruck sehr ungleich verteilt, wodurch in der Synchronie ein Gradualitätseffekt entsteht. Sodann handelt es sich bei dem Phänomen der Fixierung um einen diachronen und dynamischen und somit intrinsisch skalaren Vorgang (Mejri 2000). So kann man beobachten, dass Ausdrücke, die heute fixiert sind, in einem älteren Sprachzustand frei zusammengesetzt waren (vgl. Fournier, 1998).

Welche sind die Mechanismen, die dem Prozess der Fixierung zugrunde liegen? Unter ihnen sind zweifellos die kombinatorischen Affinitäten zwischen den Konstituenten zu nennen, deren zentrale Rolle für den Gebrauch und die Entwicklung der Sprache in den neueren Arbeiten zu Kollokationen (u. a. Butler 2003, Grossmann & Tutin, 2003, Mel'čuk 2003, Croft & Cruse 2004, Blumenthal & Hausmann 2006) aufgezeigt worden ist. Das vorliegende Projekt nimmt daher - im Rahmen eines in erster Linie quantitativen und probabilistischen Ansatzes - eine diachrone und prozessuelle Sichtweise auf die Fixierung ein. Den Ausgangspunkt unserer Forschungsarbeit bildet die in Blumenthal (2011, vgl. weiter unten) für die Untersuchung von Präpositionalphrasen dargelegte Methodik.

2.2. Eigene Vorarbeiten zum Thema

2.2.1. Arbeiten der deutschen Forschergruppe

Die Arbeiten der deutschen Forschergruppe betreffen einerseits allgemein die Methodik der statistischen und probabilistischen Analyse von Präpositionen, andererseits besondere Probleme der diachronen Betrachtung. Sie stützen sich dabei auf beträchtliche Text- und Software-Ressourcen.

2.2.1.1. P. Blumenthal (2008) verwendet statistische Methoden, die es erlauben, den Grad der Ähnlichkeit zwischen dem kombinatorischen Verhalten zweier oder mehrerer Präpositionen in einem Korpus zu ermitteln. Ausgehend vom kombinatorischen Profil (Berechnung der spezifischen Begleiter) jeder der zu vergleichenden Präpositionen gelangt man dazu, weitgehend automatisch die gemeinsame Untermenge der Begleiter sowie die getrennten Untermengen, welche die für eine einzelne Präposition charakteristischen benachbarten Elemente aufführen,

zu bestimmen. Die Ausweitung dieses Verfahrens auf die Gesamtheit der Präpositionen im Rahmen einer Berechnung durch multidimensionale Skalierung liefert eine Gesamtschau der distributionellen Gemeinsamkeiten und Unterschiede von Präpositionen und zeigt Realität und Grenzen ihrer Quasi-Synonymie in besonderen Kontexten.

2.2.1.2. Unter diachronem Gesichtspunkt wendet Blumenthal (2011) diese Methode auf zwei Korpora an; davon stammt eines aus dem 17. Jh., das andere aus der zweiten Hälfte des 20. Jh. Die Untersuchung zielt darauf ab, Frequenzverschiebungen innerhalb der Gruppe der wichtigsten Präpositionen zu erfassen, beschäftigt sich aber auch mit der teilweisen Umverteilung der semantischen und syntaktischen Rollen der Präpositionen. Es wird insbesondere gezeigt, wie im Laufe dreier Jahrhunderte einige fixierte Präpositionalphrasen untergehen, während andere sich neu bilden oder ihren semantischen oder gar textuellen Wert ändern. Die letzten Ursachen dieser Entwicklungen liegen in manchen Fällen in Bereichen des grammatischen Systems fernab von dem der Präpositionen; so bleibt die Umstrukturierung der Aktualisierung des Nomens im 17. Jh. nicht ohne Auswirkung auf den Gebrauch der Präpositionen *en* und *dans*.

Im Rahmen dieser diachronischen Fragestellung befasst sich Blumenthal (2007) mit einem ganz besonderen Problem, nämlich der Entstehung und Ausbreitung der Kombination *en* + *début* + Zeitraum (zum Beispiel *en début de semaine*) ab den 1920er Jahren. Es geht ihm dabei darum zu zeigen, was der quantitativ bemerkenswerte Erfolg dieser Konstruktion außersprachlichen Gegebenheiten, insbesondere einer neuen Art der Arbeitszeitplanung, welche auch den Soziologen aufgefallen ist, schuldet.

2.2.1.3. Auf technischer Ebene stützen sich die in Köln durchgeführten Arbeiten auf die Implementierung und Auswertung romanistischer Korpusarchive sowie auf die Entwicklung diverser textometrischer Anwendungen. Die Gesamtheit der verfügbaren Korpora liegt in lemmatisierter und morphosyntaktisch annotierter Form vor. Ein Teil der eingerichteten Samples ist ebenfalls mit Annotationen hinsichtlich syntaktischer Abhängigkeitsverhältnisse ausgestattet. Die Korpusarchive werden über den Query-Processor IMS Corpus Workbench (CQP/CWB, vgl. <http://cwb.sourceforge.net/>) verwaltet und sind mit mehreren in der Programmiersprache Perl geschriebenen Anwendungen verbunden, über welche folgende textometrische Verfahren durchgeführt werden können:

- Berechnung von Kookkurrenzstatistiken ausgehend von KWIC- Konkordanzen (vom Englischen: Keywords in Context) mittels unterschiedlicher Assoziationsmaße (log-likelihood, mutual information, z-score und exact Fisher test);
- Berechnung frequentieller Unterschiede zwischen Korpussamples;
- Berechnung multivariater Statistiken (multidimensionale Skalierung, hierarchische Partition, AFC) auf der Grundlage frequentieller und kookkurrentieller Unterschiede;
- Charakterisierung der inneren Aufteilung von Lexikogrammen mittels Stereotypie-Indices

Ein Teil der erwähnten Verfahren werden in Blumenthal/Diwersy/Mielebacher (2005), Blumenthal (2006 ; 2008 ; 2010) sowie Diwersy (2007) vorgestellt.

2.2.2. Arbeiten der französischen Forschungsgruppen

Die französische Forschergruppe vereint Linguisten und Informatiker, die seit mehreren Jahren in unterschiedlichen Konstellationen an gemeinsamen finanzierten Projekten arbeiten. Ihre jeweiligen Spezialgebiete werden im Folgenden erläutert.

Syntax und Semantik von Präpositionen: Nach seiner Doktorarbeit (2004), in der er sich den Präpositionalphrasen, welche die Präposition *en* als Kopf haben, gewidmet hatte, hat **D. Vigier**⁶ seine syntaktischen und semantischen Forschungen über PPs - insbesondere über PPs in losgelöster Stellung (2003, 2005, 2008 a,b,c; 2009) - fortgesetzt. **D. Leeman** befasst sich seit der Mitte der 80er Jahre mit dem syntaktischen Verhalten von Konstruktionen im Bereich von indirekten Objekt und Umstandsangaben. Diese traditionellen Bezeichnungen tragen allerdings den tatsächlichen Eigenschaften der grammatischen Erscheinungen nur unzureichend Rechnung. Deren Analyse muss sich auf das stützen, was beobachtbar und manipulierbar ist, d.h. auf materialisierte Formen, die den üblichen distributionellen und syntaktischen Tests unterzogen werden (Tilgungs-, Ersetzungs-, Hinzufügungs-, Hervorhebungs-, Fragetest usw.) und dadurch als Interpretationsgrundlage dienen können (Leeman 1999). **B. Fagard** hat sich in seiner Doktorarbeit (2006) der Untersuchung der semantischen Entwicklung von Präpositionen gewidmet und seitdem weiter in diesem Bereich geforscht, wobei er stets diachronen Korpusuntersuchungen im Rahmen der Grammatikalisierungstheorie den Vorzug gegeben hat. **L. Gardelle** hat die Kombinatorik der Präposition *of*, und im Einzelnen die Konkurrenz von *thereof* und *of it*, untersucht.

Diachronie: Neben B. Fagard sind auch **C. Guillot** und **A. Lavrentiev** Spezialisten für mittelalterliches Französisch und die Diachronie des Französischen. Sie entwickeln gemeinsam die Base de Français Médiéval (<http://bfm.ens-lyon.fr/>) und das Corpus de très ancien français CoRPTeF (<http://corptef.ens-lyon.fr/>). Sie nehmen ebenfalls am DFG/ANR-Projekt Syntactic Reference Corpus of Medieval French (<https://listes.cru.fr/wiki/srcmf/>) teil. **M.-L. Demonet** hat als Mitglied des IUF und Spezialistin für das 16. Jh. über Rabelais, Montaigne, Pasquier (kritische und elektronische Ausgaben, Tagungsberichte, Monographien) veröffentlicht. Sie hat zwei Internetseiten geschaffen, deren letztere unveröffentlichte Texte aus der Renaissance enthält (<http://www.bvh.univ-tours.fr/Epistemon>); zudem ist sie für das Projekt „Bibliothèques virtuelles des humanistes“ verantwortlich. **S. Rémi-Giraud** ist Spezialistin für Syntax und Semantik des Französischen und hat sich in ihrer Habilitationsschrift einer semantischen Untersuchung des Wortes *air* ('Erscheinung') im 17. und 20. Jh. gewidmet. Die Studie über das 17. Jh. ist hauptsächlich anhand der Correspondance von Mme de Sévigné durchgeführt und anschließend auf über zwanzig weitere Autoren erweitert worden.

Diskursrahmen: **L. Sarda** und **M. Charolles** forschen seit mehreren Jahren aus einer diskursiven Perspektive über Adverbiale – darunter auch PPs. M. Charolles hat sich in zahlreichen Veröffentlichungen (Artikel, Herausgabe von Zeitschriftennummern) den rahmenbildenden Adverbialen gewidmet und gezeigt, dass diese eine neue Gruppe von diskursiven Kohäsionsmarkern bilden, in gleicher Weise wie die Konnektoren und anaphorischen Ausdrücke. Gemeinsam haben sie ein ANR-Projekt zu Diskursrahmen mit räumlicher Bedeutung (Spatial Frames Adverbials) geleitet (2006-2009, <http://www.lattice.cnrs.fr/Resume-Projet>), an dem auch B. Fagard und D. Vigier teilgenommen haben.

Werkzeuge und Textometrie: **S. Heiden** und **B. Pincemin** sind entscheidend an textometrischer Forschung beteiligt und steuern die Entwicklung der offenen Software-Plattform TXM, welche frei im Netz verfügbar ist. Diese Plattform knüpft an die textometrische Tradition an (Konkordanzen, Spezifitätsberechnungen, Kookurrenzen, Faktorenanalyse der Entsprechungen, usw.) und bereichert diese, indem sie sie an die aktuellen (strukturierten und annotierten) Korpora anpasst. Sie beinhaltet die internationalen Standards (XML, TEI) und besitzt einen modularen

⁶ <http://icar.univ-lyon2.fr/membres/dvigier/publications.htm>

Aufbau, der es ihr ermöglicht, leistungsstarke Komponenten (CQP Korpusindexierungs- und -suchmaschine) zu nutzen.

3. Ziele und Arbeitsprogramm

3.1. Ziele

3.1.1. Wissenschaftliche Ziele

a) Vertiefung unserer Kenntnisse zur Kombinatorik der französischen Präpositionen von den Anfängen bis zur Gegenwart

Für den Zeitraum vom 9. bis 20. Jh. werden wir mit statistischen Methoden die diachrone Entwicklung des „kombinatorischen Profils“ (Blumenthal 2008) jeder einzelnen der untersuchten Präpositionen, vornehmlich durch Bestimmung ihrer spezifischen Begleiter⁷, zu Tage fördern. Durch eine multidimensionale Analyse (Faktorenanalyse und multidimensionale Skalierung) werden kontextuelle Ähnlichkeiten zwischen den unterschiedlichen untersuchten Präpositionen ermittelt. Dank der ins Korpus eingeführten Annotierungen werden wir über Informationen zur Morphosyntax sowie zu den Diskursgattungen der Korpus Texte verfügen. All diese Informationen stellen Variablen dar, die es ermöglichen werden, den Ansatz des kombinatorischen Profils von Präpositionen zu verfeinern. Die dem vorliegenden Antrag beigefügten Anhänge (darunter Blumenthal 2012) veranschaulichen die angewandte Methode und die herangezogenen Werkzeuge.

b) Erfassung der Emergenz, der Umgestaltung und des Schwindens nicht-freier Wortverbindungen

Die Verwendung statistischer Methoden bei der Untersuchung signifikant häufiger Wortverbindungen wird mehr oder weniger feste (nicht-freie) Präpositionalgruppen zu Tage fördern. Unser Projekt zielt u. a. auf eine genaue diachrone Untersuchung solcher Kombinationen ab. Es geht dabei um Entstehung und Entwicklung fester Wendungen, um ihren möglichen Bedeutungswandel, die Änderung ihrer spezifischen Umgebung im Satz und ihrer Stellung usw. Es soll versucht werden, eine Beziehung zwischen solchen Erscheinungen und weitergehenden Umstrukturierungen des Sprachsystems zu erfassen, die im Rahmen der Grammatikalisierungstheorie interpretiert werden können (vgl. u. a. Heine, Claudi & Hünnemeyer (1991), Hopper & Traugott (Hgg.) (1993), Heine & Kuteva 2002). Hauptgegenstände der Untersuchung sind Grammatikalisierungsketten (u. a. Heine, Claudi & Hünnemeyer (ibid), Traugott & König (1991)), die Herausbildung fester Wendungen und die Stellung der Präpositionalphrase im Satz.

c) Hypothesenbildung zu Zusammenhängen zwischen Frequenzschwankungen in der Präpositional-Kombinatorik, der Suche nach semantischen oder diskursiven Mehrwerten und den Ausdrucksbedürfnissen innerhalb einer Gesellschaft

Das im vorigen Punkt skizzierte Programm, welches quantitative Beschreibung mit qualitativer Analyse in Zusammenhang bringt, soll über die strukturelle Dimension hinaus ausgeweitet werden; im Anschluss an Coseriu (1974: 246) werden wir den

⁷ Häufigste Nomen und andere Wortarten im linken und/oder rechten Umfeld, gemessen an der erwartbaren absoluten Häufigkeit.

Ausdruckbedürfnissen nachgehen, die sich in der Gesellschaft einer bestimmten Epoche ausmachen lassen und die ursächlich für die soeben unter a) angesprochenen kombinatorischen und strukturellen Veränderungen sein können. In diesem Zusammenhang gewinnt der Begriff des „Mehrerts“, welcher einer bestimmten, vor allem nicht-freien präpositionalen Kombination innewohnt, erst seine volle Bedeutung. Eine Argumentation dieser Art findet sich in Blumenthal 2012 bezüglich einiger Verbindungen von *en* (zum Beispiel *en ville*).

d) Hypothesenbildung zu Zusammenhängen zwischen einigen der beobachteten Umstrukturierungen und gesellschaftlichen Entwicklungen

Ein weiterer origineller Ansatz des Projektes liegt im Versuch, Entwicklungen im Bereich der festen Wendungen, die als « faits sociaux » im Sinne Durkheims gelten dürfen, mit anderen gesellschaftlich relevanten Tatsachen zu verbinden (s. oben 2.2.1.2.). Wir sind der Auffassung, dass sich nicht alle grammatischen Entwicklungen auf sprachinhärente Prinzipien wie die Grammatikalisierung zurückführen lassen. In diesem Sinne soll versucht werden, die traditionelle Opposition zwischen externer und interner Sprachgeschichte zu überwinden und nach einem dritten Weg zu suchen, der sich an aus den beiden entgegen gesetzten Bedingungsbeziehungen stammenden Faktoren orientiert (vgl. Blumenthal 2003, z. B. zur Verbindung *en début de semaine*, die in Zusammenhang mit einem bestimmten Typ von Arbeitsorganisation in Industriegesellschaften steht).

e) Anwendung des Begriffes des kombinatorischen Profils (Blumenthal 2006) auf Kookkurrenz-schemata in größeren Texteinheiten

Charolles (1997) hat gezeigt, dass vor allem räumliche und zeitliche adverbiale Präpositionalphrasen am Satzbeginn häufig Funktionen der Textstrukturierung übernehmen. Beispiel:

[Dans les zones rurales, les personnes de plus de 25 ans sont avantagées par rapport aux jeunes, car elles entrent bien moins fréquemment en chômage]. [Dans les zones urbaines c'est l'inverse : les jeunes entrent relativement moins souvent en chômage ou en sortent plus facilement comme à Lille, Lens et Douai]. Par contre, les rythmes de sortie du chômage ne sont pas aussi différenciées ... (Les marchés locaux du travail dans le Nord-Pas-de-Calais, Economie et statistique, 182, p. 57).

Die Funktionsweise der Adverbiale in solchen Kontexten ist für das Französische in jüngster Zeit wiederholt untersucht worden (u. a. Charolles & Péry-Woodley (eds) 2009, Charolles et Prévost (eds) 2003, Vigier & Terran (eds) 2005) sowie das ANR-Projekt Blanc 2006-2009 « Spatial Framing Adverbials » in UMR LaTITCe). Diese Studien haben im Anschluss an Thompson (1985) und Virtanen (1992a und 1992b) gezeigt, dass nur vorangestellten Präpositionalphrasen (zu Beginn des Satzes und vor allem des Abschnitts) eine solche textstrukturierende Kraft zukommt.

Das erwähnte Modell der Textstrukturierung lässt sich in deskriptiver Hinsicht verbinden mit Hoeyes im Rahmen der *Lexical Priming Theory* entwickelten Modell, zumal im Hinblick auf textsemantische Assoziationen, die auf Kookkurrenzketten beruhen (Hoey 2005:122ss). In diesem Sinne schlagen wir vor, den Anwendungsbereich des Begriffes des kombinatorischen Profils auf Kookkurrenz-schemata in größeren Texteinheiten auszudehnen. Nach Hoey tragen Kollokationen (im Sinne von Firth) zur Strukturierung von Texten bei, da Wörter und Syntagmen Teile von Assoziationsnetzen sind, die auf dem Erfahrungswissen des

Sprechers sowie auf den üblichen kontextuellen Strukturen innerhalb von Gattungen und Diskurstraditionen beruhen.

Für die Beschreibung dieser Strukturen sieht Hoey (2005 :13) ein Modell vor, das zwischen zwei Typen des Priming unterscheidet : lexikalische Kombinatorik im engen Kontext und semantische Wechselwirkungen zwischen Lexemen auf der Textebene unter dem Einfluss ihrer spezifischen Stellung im Diskurs („diskursive Kolligationen“).

Im vorliegenden Projekt werden wir uns in diesem Zusammenhang mit folgenden Fragen beschäftigen :

- ✓ Zieht eine textstrukturierende Präpositionalphrase in Anfangsstellung normalerweise das Auftreten ähnlicher Konstruktionen im weiteren Text nach sich?
- ✓ Werden derartige textstrukturierende Phrasen typischerweise in einem vorausgehenden Satz angekündigt? Dies ließe auf eine längerfristige Textplanung schließen.
- ✓ Welche Rolle kommt den Texttypen hinsichtlich der Auftretenshäufigkeit textstrukturierender Präpositionalphrasen zu?

Unsere Untersuchung wird sich mit Rücksicht auf die verfügbaren Texte auf die Periode vom 17. bis zum 20. Jh. beschränken⁸.

f) Zusammenstellung eines ausreichenden Korpus für die Gesamtperiode (9. bis 20. Jh.)

Die Zusammenstellung eines adäquaten Korpus ist eine wesentliche Bedingung für die Durchführung des Vorhabens. Wir werden hierfür in größtmöglichem Umfang auf bereits bei beiden Partnern vorhandene Korpora zurückgreifen, sie noch weiter ausbauen und annotieren.

In chronologischer Hinsicht wird das Gesamtkorpus die Zeit von den ersten französischen Texten (9. Jh.) bis zum Ende des 20. Jhs. abdecken. Der mittelalterliche Teil (9.-15. Jh.) wird aus Texten der *Base de Français Médiéval* (BFM) bestehen, aber auch auf Texte aus dem *Nouveau Corpus d'Amsterdam* (NCA) zurückgreifen. Die Texte zum 16. Jh. stammen aus der Datenbank *Epistemon*⁹ unseres Projektpartners *Bibliothèques Virtuelles Humanistes* sowie aus Textbeiträgen des ATILF. Für die folgende Zeit benutzen wir die *Kölner Romanistische Datenbank*, die durch UNIVC¹⁰, ARTFL¹¹, FRANTEXT und gegebenenfalls durch frei zugängliche Online-Textressourcen (CNRTL, Gallica, Wikisource, Gutenberg) ergänzt wird. Was die Anzahl der Wörter anbelangt, werden wir einen einzigen Block mit 500 000 Wörtern für die ältesten Texte (9.-12. Jh.), die ja zahlenmäßig gering sind, zusammenstellen. Für die Zeit vom 13. bis zum 15. Jh. werden wir über Blöcke von 500 000 Wörtern pro Jahrhundert verfügen. Für die Zeit vom 16. bis zum 20. Jh. wird jeder Abschnitt von 50 Jahren durch ungefähr 5 Millionen Wörter vertreten sein.

Wir werden versuchen, im Gesamtkorpus ein Gleichgewicht zwischen verschiedenen Textsorten und Gattungen herzustellen. Die genaue Verteilung wird zu

⁸ Als Grundlage für eine derartige Studie ist in der Tat ein sehr umfangreiches Textvolumen erforderlich. Im übrigen entwickelt sich der französische Satzbau erst im 15. Jh. zu einer Struktur, in der die Anfangsstellung der Präpositionalphrase textlinguistische Relevanz gewinnen kann (vgl. C. Marchello, 1999 : 58 ss ; 2006 : 133).

⁹ <http://www.bvh.univ-tours.fr/Epistemon/index.asp>. 44 annotierte Texte dieser Datenbank sind bereits auf die Plattform TXM hochgeladen worden.

¹⁰ (<http://marriage.uvic.ca/toc.htm?sort=date asc>)

¹¹ (<http://www.lib.uchicago.edu/efts/ARTFL/projects/BibB1/>)

Beginn des Projektes nach den in der Base de Français Médiéval verwendeten Kriterien durchgeführt. Typologische und quantitative Gesichtspunkte können durch Qualitätsmerkmale der Informatisierung und Annotierung ergänzt werden.

Frei zugänglichen Texten geben wir grundsätzlich den Vorzug. Allerdings wird der Zugriff auf private Textressourcen bisweilen unvermeidlich sein. Alle Texte werden entsprechend den Empfehlungen der TEI automatisch verarbeitet und tokenisiert, getaggt sowie mit dem TreeTagger oder TnT lemmatisiert. Die bereits manuell vorgetaggtten 30 mittelalterlichen Texte (ca. 500.000 Wörter) dienen als Trainingskorpus im Hinblick auf den Einsatz der informatischen Werkzeuge.

g) Entwicklung von Lemmatisierungsprogrammen für das Französische für alle Perioden des Korpus

Das Projekt wird je nach den Erfordernissen der verschiedenen Perioden des Korpus ein oder mehrere Lemmatisierungsprogramme für das Französische entwickeln. Dazu wird es Trainingssoftwares zur Erstellung linguistischer Modelle wie TreeTagger (Schmid, 1995) oder TnT (Brants, 2000)¹² verwenden. Diese linguistischen Modelle werden für die Lemmatisierung und Annotation des PRESTO-Korpus optimiert und online veröffentlicht werden, um für andere Forschungsprojekte zur Verfügung zu stehen.

h) Schaffung einer einheitlichen Software-Plattform für die distributionelle und kombinatorische Analyse

Das anvisierte einheitliche Analyseverfahren verbindet die textometrische Herangehensweise mit Praktiken der linguistischen Korpusanalyse. Sie kombiniert die in Lyon und Köln entwickelten informatischen Ressourcen und Fähigkeiten. Beiden Arbeitsgruppen sind bereits mehrere technologische Grundentscheidungen gemeinsam: Sie verwenden die Software CWB/CQP für die Korpusindizierung und die R-Bibliothek für die statistische Analyse. Sie verwenden die gleichen Berechnungsverfahren für Konkordanzen, Kookkurrenzen und multivariate Statistiken. Die Zusammenarbeit wird diese technologischen Gemeinsamkeiten festigen und vertiefen, und zwar u.a. in folgenden Bereichen:

- ✓ noch genauere Berechnung der kombinatorischen Profile unter Einbeziehung syntaktischer Dependenzrelationen;
- ✓ nuanciertere Interpretation der Diagramme der multidimensionalen Skalierung durch Rückgriff auf die Ausgangstexte;
- ✓ Einfügung aller Analysemodule in eine open-source-Umgebung und eine umfassende, zweifellos für eine größere Forschergemeinschaft attraktive Plattform (Windows, Linux, Macintosh). Diese entsteht im Rahmen der Lyoner Textometrie-Plattform TXM und bildet eine Voraussetzung für die Entwicklung neuer Funktionalitäten im Hinblick auf die besonderen Projektziele im Bereich der Diachronie und der festen Wendungen.

Die Webportalversion der Plattform TXM wird es schließlich ermöglichen, die im Projekt erzeugten Korpora online zu veröffentlichen.

3.1.2. Zu erwartende Synergien

Dieses Projekt stellt eine weitere Etappe in der Zusammenarbeit zwischen beiden Forschergruppen dar, die mit der Planung einer der Präposition *en* gewidmeten

¹² Vgl. Schmid (1995) sowie Brants (2000).

Nummer der Zeitschrift *Langue Française*¹³ (mit D. Vigier als Herausgeber und unter Beteiligung von P. Blumenthal) begann.

Die angestrebte Synergie wird sich aus der Komplementarität der jeweiligen Forschungsbereiche und aus der grundlegenden Konvergenz der theoretischen, methodischen und technischen Orientierungen der Projektarbeit ergeben. Eines der bevorzugten Gebiete dieser Komplementarität wird die Untersuchung der Struktur derselben Texte unter den unterschiedlichen Blickwinkeln der beiden Methoden sein, von denen eine von unserer „Stammzelle“ (d.h. von der Präpositionalphrase samt ihrem kombinatorischen Profil) und die andere von einer distributionellen Sicht auf den Gesamttext ausgeht.

3.1.2.1. Im linguistischen Bereich arbeiten einige Mitglieder der Forschergruppen in Lyon und Paris seit mehreren Jahren über verschiedene französische Präpositionen, und zwar aus synchroner (Vigier, Sarda) und aus diachroner (Fagard) Sicht. Ihre Untersuchungen sind korpusbasiert, verwenden aber nicht statistische Methoden und informatische Werkzeuge von der Art, wie sie im Projekt Presto vorgesehen sind. Die Kölner Gruppe hat ebenfalls, wenn auch weniger regelmäßig, über Präpositionen gearbeitet (Blumenthal 2008 und 2011). Sie verfügt hingegen über eine große Erfahrung in der Untersuchung von kombinatorischen Profilen zu Basiswörtern, die verschiedenen Wortarten angehören. Das Projekt wird den französischen und deutschen Forschern die Möglichkeit geben, ihre jeweiligen Fähigkeiten und Erfahrungen in produktiver Weise miteinander zu verbinden.

Im Übrigen haben mehrere Mitglieder der beiden französischen Forschergruppen über Diskursstrukturierung gearbeitet, so Charolles, der seit einem Jahrzehnt verschiedene Forschungsvorhaben zu dieser Thematik initiiert hat. Das hier beantragte Projekt wird die Voraussetzung für eine methodische, in konkreter Korpusarbeit verankerte Verbindung zwischen den vorgenannten Ansätzen (vgl. auch Charolles, 1997) und der auf ähnliche Ziele ausgerichteten Theorie des *lexical priming* (Hoey 2005) schaffen. Empirischer Ausgangspunkt dieser gemeinsamen Orientierung ist die textstrukturierende Funktion vorangestellter Präpositionalphrasen.

3.1.2.2. Die gemeinsame Arbeit im Bereich der Textstatistik wird zu einer Verbindung von Korpuslinguistik und Textometrie und zu einer gegenseitigen methodischen Ergänzung und Bereicherung dieser Forschungszweige führen.

3.1.3. Veröffentlichung der Ergebnisse

3.1.1.1 Ressourcen

Wir unterscheiden hinsichtlich der Veröffentlichung der Ergebnisse zwischen Textkorpora und Software zur automatischen Korpusanalyse.

Hinsichtlich der Korpusressourcen unterscheiden wir:

- Annotierte Textkorpora (lemmatisiert, geparst, usw.)
- Trainingskorpora zur Erstellung von Werkzeugen zur automatischen Korpusannotierung
- Lexikodateien (mit Wortformen, morpho-syntaktischen Angaben, Lemmaangaben etc.) zur Erstellung von Annotierungswerkzeugen.

Hinsichtlich der Analysesoftware unterscheiden wir:

¹³ Dieses Projekt wird derzeit durch das wissenschaftliche Komitee von *Langue Française* begutachtet.

- Annotierungswerkzeuge und damit verbundene Ressourcen (Parameterdateien, Parsing-Modelle, etc.)
- Software zur automatischen Korpusauswertung
- (on-line zugängliche) Plattformen zum Herunterladen und Abrufen von Korpora.

Bezüglich der Veröffentlichung der Korpora (a) werden wir uns an folgenden Prinzipien orientieren¹⁴:

Wir unterscheiden zwei Hauptarten der Veröffentlichung: herunterladbar (offenes Korpus) und über eine Plattform oder ein Portal zugänglich (beschränktes Korpus).

Unser Anliegen ist es, eine möglichst große Anzahl von Korpora gemäß einer noch zu definierenden offenen Lizenz (des Typs Creative Commons) in herunterladbarer Form zu veröffentlichen, wohl wissend, dass es aufgrund bestimmter Verpflichtungen gegenüber Partnern, welche die Quelldokumente für bestimmte Korpora liefern, nicht möglich sein wird, alles frei zugänglich zu machen. Für Ressourcen, die derartigen Beschränkungen unterliegen, wird eine Veröffentlichung über das Projekt-Portal gegen obligatorische Anmeldung nach Abschluss einer Zugangsvereinbarung angeboten; dabei werden wir uns auf die Vereinbarung stützen, die bereits für den Zugang zur *Base du Français Médiéval* (BFM) verwendet wird und in Zusammenarbeit mit der Rechtsabteilung der ENS und der Delegation Rhône-Auvergne des CNRS (siehe Anhang X) ausgearbeitet wurde.

Innerhalb eines solchen Portals können nach den jeweiligen Beschränkungen, denen eine Ressource unterliegt, verschiedene Zugangsmodalitäten vorgesehen werden¹⁵, was sich anhand des Veröffentlichungsportals der BFM¹⁶ (<http://txm.risc.cnrs.fr/bfm>) nach gegenwärtigem Stand wie folgt illustrieren lässt:

- I. 42 Texte im Umfang von insgesamt 1,5 Millionen Wörtern sind gemäß einer Creative Commons Lizenz vom Typ BY-NC¹⁷ zum Herunterladen und für jedwede Korpusanalyse, die sich mithilfe der Portalwerkzeuge durchführen lässt, verfügbar;
- II. 25 Texte im Umfang von insgesamt 0,9 Millionen Wörtern sind für jedwede Korpusanalyse, die sich mithilfe der Portalwerkzeuge durchführen lässt, zugänglich, jedoch nicht herunterladbar;
- III. 18 Texte im Umfang von insgesamt 0,6 Millionen Wörtern sind für jedwede Korpusanalyse, die sich mithilfe der Portalwerkzeuge durchführen lässt, zugänglich, wobei die Größe der angezeigten Konkordanz-Kontexte bei Suchanfragen dauerhaft begrenzt ist und es nicht möglich ist, die Ausgabe eines Textes Seite für Seite zu lesen (das Herunterladen ist dabei verboten).

Bezüglich der Trainingskorpora und Lexika (b) und (c) ist es unser Anliegen, eine größtmögliche Anzahl gemäß einer noch zu definierenden offenen Lizenz (des Typs Creative Commons) in herunterladbarer Form bereitzustellen. Trainingskorpora, die (aus jedweden rechtlichen Gründen wie dem Schutz geistigen Eigentums usw.) Zugangsbeschränkungen unterliegen, werden nicht durch das Projekt Presto veröffentlicht. Als Orientierungsmaßstab dienen uns dabei grundsätzlich die hinsichtlich des für das Altfranzösische entwickelten Taggers CATTEX getroffenen Regelungen, die ein Herunterladen der entsprechenden Ressourcen wie Trainingskorpus und Vollformenlexikon gemäß einer offenen Lizenz vom Typ Creative Commons oder LGPLLR¹⁸ ermöglichen.

Bezüglich der Annotierungswerkzeuge und der damit verbundenen Parameter- und Modelldateien, die im Rahmen des Projekts erstellt oder verbessert werden (d), ist es unser Anliegen, eine größtmögliche Anzahl gemäß einer offenen noch zu definierenden Lizenz (des Typs GPL V3 oder LGPLLR je nach Fall) in herunterladbarer Form zu veröffentlichen.

¹⁴ Als Ausgangspunkt dienen uns hierbei die für die Veröffentlichung der *Base du Français Médiéval* (BFM) zu Grunde gelegten Projektrichtlinien.

¹⁵ Für die Erstellung von Veröffentlichungslizenzen in Verbindung mit der Rechtsabteilung der ENS wird der Rat spezialisierter Juristen erforderlich sein (siehe Budget).

¹⁶ Mit ihrem Textbestand, welcher den Zeitraum vom 9. bis zum 16. Jh. abdeckt, wird die BFM den Hauptteil des PRESTO-Korpus für die Periode des Altfranzösischen ausmachen.

¹⁷ <http://creativecommons.org/licenses/by-nc/2.0/fr>

¹⁸ <http://infolingu.univ-mlv.fr/DonneesLinguistiques/Lexiques-Grammaires/lgpllr.html>

Bezüglich der verschiedenen Anwendungen zur automatischen Korpusanalyse, die durch das Projekt entwickelt oder verbessert werden (e), ist es unser Anliegen, eine größtmögliche Anzahl gemäß einer offenen noch zu definierenden Lizenz (des Typs GPL V3) in herunterladbarer Form kostenlos zur Verfügung zu stellen.

Wir orientieren uns dabei an den für die online und lokal einsetzbare Korpussoftware TXM geltenden Bereitstellungsregelungen.

Schließlich soll nach Maßgabe der vorhandenen Beherbergungsinfrastruktur für die Laufzeit des Projektes (und – in dem Maße, wie es die durch die Institutionen der Projektpartner bereitstellbaren technischen Ressourcen zulassen – auch über seine Beendigung hinaus) über eine entsprechende Veröffentlichungs- und Abfrageplattform ein dauerhafter Zugriff auf das durch Presto ins Netz gestellte Korpus ermöglicht werden.

3.1.3.2. Veranstaltungen und Veröffentlichungen

Im Jahre 2014 werden wir in Köln eine eintägige Arbeitstagung veranstalten, zu der mehrere internationale Forschergruppen, die über diesen Themenbereich arbeiten, sowie die am Projekt beteiligten Berater eingeladen werden. Diese Tagung wird den Projektmitarbeitern Gelegenheit geben, der wissenschaftlichen Öffentlichkeit erste Ergebnisse ihrer Forschungen vorzustellen.

Nach Abschluss des Projektes wird in Lyon im Jahre 2015 ein Kolloquium mit internationaler Ausschreibung stattfinden.

Eine größere gemeinsame Veröffentlichung wird über die im Projekt Presto erreichten Ergebnisse berichten.

Die Ergebnisse der linguistischen Analyse werden auf der Website des Projekts über eine frei zugängliche Abfrageschnittstelle sowie in Form herunterladbarer SQL-Quelldateien gemäß einer Lizenz des Typs Creative Commons der Öffentlichkeit zur Verfügung gestellt.

Die im Rahmen der TXM-Plattform erstellten Programme werden unter GPL-Lizenz (v. 3) auf SourceForge¹⁹ (Webseite für freie Software) als quelloffene Software zugänglich sein.

3.2 Arbeitsprogramm, Methoden, Zeitplan

Die Planung und Verteilung der Arbeitspakete ist entsprechend den jeweiligen Kompetenzen der Mitarbeiter erfolgt: In jeder Forschergruppe besitzt ein Forscher eine doppelte Kompetenz in Informatik und Linguistik (S. Heiden in Lyon, S. Diwersy in Köln), während die anderen Forscher entweder Linguisten oder Informatiker sind.

3.2.1 Arbeitsprogramm und Methoden

3.2.1.1 Programm und Methodik der sprachwissenschaftlichen Untersuchung

Arbeitspaket AP 1: Pilotstudie zu der im Zeitraum vom 16. bis 20. Jh. zu beobachtenden Entwicklung des systeminternen Werts der Präpositionen *en* und *dans* auf Grundlage der bereits vorliegenden Korpora

Teilpaket TP 1: Untersuchung des Verhältnisses der beiden Präpositionen im vorklassischen Französisch (differentielle Kookkurrenzanalyse auf der Grundlage kombinatorischer Profile und multivariater Auswertungstechniken wie der MDS u.a.)

Teilpaket TP 2: Untersuchung des Verhältnisses der beiden Präpositionen im klassischen Französisch

¹⁹ Einige wenige Implementierungen von im Projekt benutzten Algorithmen werden möglicherweise nicht frei zugänglich gemacht. Die Art der möglichen Beschränkung wird von dem Projektpartner verantwortet, der diese verlangt.

Teilpaket TP 3: Analyse der Verwendung von *en* und *dans* in der *Encyclopédie* von Diderot und d'Alembert (zweite Hälfte des 18. Jhs.)

Teilpaket TP 4: Untersuchung der Entwicklung bis zum Zeitraum um 1880

Teilpaket TP 5: Analyse der ab dem Ende des 19. Jhs. zu verzeichnenden Tendenzen

Teilpaket TP 6: Gesamtauswertung der Ergebnisse hinsichtlich der in Abschnitt 3.1 erwähnten Aspekte (Bildung fester Wendungen, Grammatikalisierung, Lexikalisierung, vergleichende Betrachtung anderer grammatischer Subsysteme (z.B. Substantivdetermination) textuell-argumentative Rahmensetzung durch bestimmte Präpositionalsyntagmen wie *en* + NP / *dans* + NP)

Arbeitspaket AP 2: Anwendung der in AP 1 entwickelten Methodik auf ältere Sprachzustände (vom 9. bis zum Ende des 16. Jhs.) unter Einbeziehung der 10 bis 15 wichtigsten Präpositionen

Teilpaket TP 1: Untersuchung der Entwicklungen vom 9. bis 12. Jh.

Teilpaket TP 2: Analyse des Präpositionalgebrauchs im Werk von Chrétien de Troyes

Teilpaket TP 3: Analyse des Präpositionalgebrauchs in der narrativen Prosa des Mittelfranzösischen

Teilpaket TP 4: Analyse des Präpositionalgebrauchs in den explikativen und argumentativen Texten des 16. Jhs.

Teilpaket TP 5: Synoptische Beschreibung des Präpositionalsystems im Hinblick auf die typologische Restrukturierung im Alt- und Mittelfranzösischen und die Ausdifferenzierung verschiedener Textgattungen

Arbeitspaket AP 3: Erweiterte Analyse der Entwicklung des Präpositionalsystems vom klassischen Französisch bis heute

Teilpaket TP 1: Beschreibung der Entwicklungen unter Einbeziehung des historischen Kontexts; Einteilung des Datenbestands in jeweils zwei, gemäß aufeinander folgender Zeitabschnitte gegliederte Samples pro Jahrhundert und entsprechende Bestimmung der Ähnlichkeit des kombinatorischen Verhaltens der einbezogenen Präpositionen anhand multivariater Verfahren (MDS u.a.); diachronische Analyse der betreffenden synchronen Schnitte.

Teilpaket TP 2: Untersuchung des Texteinbettungsprofils verschiedener durch lexikalisch-syntaktische Musterbildung geprägter Präpositionalsyntagmen und ihrer diskursiven Funktion (thematisch-argumentative Rahmensetzung wie im Falle von frz. *en politique*).

Teilpaket TP 3: Erarbeitung eines Sammelbandes, der eine Synthese der im Rahmen des Projekts gewonnenen Erkenntnisse bietet.

3.2.1.2 Arbeitsprogramm und Methoden hinsichtlich der Erstellung, Verarbeitung und Veröffentlichung der projektbezogenen Sprachdaten

Arbeitspaket AP 1: Zusammenstellung, Vervollständigung und formatbezogene Normalisierung des zu verwendenden Korpusmaterials

Bereitstellung und Beschaffung elektronisch verfügbaren Textmaterials mit der Maßgabe, für jede der zu untersuchenden Zeitabschnitte einen im Hinblick auf Umfang und Textsorten ausgewogenen Korpusbestand zu gewährleisten.

Gleichzeitige Programmierung von Werkzeugen, die eine Konvertierung anderweitig formatierter Dokumente in ein Unicode- und TEI-konformes, hinsichtlich des Einsatzes der Korpusplattform TXM standardisiertes Format ermöglichen.

Arbeitspaket AP 2: Austausch und Abstimmung über Kodierungs- und Annotationskonventionen

Abprache und Festlegung der Vorgaben im Hinblick auf:

- die morphosyntaktische Annotierung (unter Berücksichtigung bestehender Tagset-Formate wie STTS, Multext, CATTEX u.a.)
- die Lemmatisierung (Aufbau des Formenlexikons)
- die Kodierung der Satzpositionen von Präpositionalsyntagmen
- die Annotierung von Abhängigkeitsrelationen zwischen Präpositionen und den entsprechenden syntaktischen Köpfen
- die Kodierung der Position von Präpositionalsyntagmen innerhalb größerer Texteinheiten
- die Kodierung von Genre- und textsortenspezifischen Angaben

Arbeitspaket AP 3 : Satz- und Wortsegmentierung, morphosyntaktische Annotation und Lemmatisierung

Automatische Korpusannotierung hinsichtlich folgender morpho-lexikalischer Angaben :

- Teilpaket TP 1: Satzsegmentierung und Tokenisierung (graphische Wortformen)
- Teilpaket TP 2: morphosyntaktische Annotation (PoS-Tagging)
- Teilpaket TP 3: Lemmatisierung

Die betreffenden Arbeitsschritte sollen anhand von Scripten durchgeführt werden, die die Einbindung und Anwendung verschiedener lern- oder regelbasierter Taggingalgorithmen, wie sie beispielsweise durch die Annotationsframeworks und -programme OpenNLP, Unitex oder TreeTagger implementiert sind, ermöglichen. Hierbei kann auf die jeweiligen Vorarbeiten der Partner bei der Erstellung und Nutzung entsprechender Werkzeuge zurückgegriffen werden.

Arbeitspaket AP 4 : Syntaktische Korpusannotation

Automatische Korpusannotierung hinsichtlich folgender syntaktischer Angaben :

- Teilpaket TP 1: Annotation syntaktischer Abhängigkeitsrelationen
- Teilpaket TP 2: Annotation syntaktischer Konstituenten

Die betreffenden Arbeitsschritte sollen anhand von Scripten durchgeführt werden, die die Einbindung verschiedener lern- oder regelbasierter Parsingalgorithmen ermöglichen, wie sie beispielsweise durch entsprechende Softwareanwendungen wie MATE Tools oder LinguaStream implementiert werden.

Arbeitspaket AP 5: Segmentierung und positionsbezogene Kodierung von satzübergreifenden Texteinheiten (Kapitel, Absätze)

Um Kookkurrenzberechnungen oberhalb der Satzebene zu ermöglichen, soll anhand entsprechender Skripte eine Abgrenzung der identifizierbaren Texteinheiten (wie beispielsweise Kapitel, Absätze u.a.) vorgenommen werden, die daraufhin mittels der über die Korpusplattform TXM aufrufbaren Dokumentansicht semi-automatisch

überprüft werden kann. Anschließend sollen die Positionen der Textobjekte im Verhältnis zum Anfang, zur Mitte oder zum Ende der jeweils nächst höheren Einheit automatisch kodiert werden.

Allgemeine Anmerkungen zu den Arbeitspaketen AP 1 bis AP 5 :

- IV. die Verarbeitung der Korpusdokumente wird unter Verwendung entsprechender Groovy- und XSLT-Scripte im Rahmen der Plattform TXM durchgeführt.
- V. Im Zuge der automatischen Annotierung kann es bedingt durch das zu verarbeitende Textmaterial zu abweichenden Ergebnissen kommen. Im Rahmen der jeweiligen Arbeitspakete gilt es daher, die Qualität der generierten Annotationen zu überprüfen und gegebenenfalls Texte mit einer allzu hohen Fehlerrate aus dem endgültigen Korpusbestand zu entfernen.

Arbeitspaket AP 6: Bereitstellung von Werkzeugen zur textometrischen Korpusauswertung

Die Korpusplattform TXM bündelt zwei wesentliche technische Komponenten, die gegenwärtig bereits von beiden Partnern gleichermaßen genutzt werden :

- die Korpusmaschine CWB-CQP, die eine effiziente Verwaltung von Textsammlungen großen Umfangs (wie etwa der vorgesehenen Arbeitskorpora mit über 50 Mio. Wortform-Token) ermöglicht.
- die Statistiksoftware R, die für die Durchführung komplexer Kookkurrenzberechnungen und den Einsatz multivariater Analysetechniken herangezogen werden kann.

Die im Zuge der Umsetzung der einzelnen Projektvorhaben jeweils notwendige Anpassung der TXM-Anwendungen soll im Rahmen der folgenden Teilpakete realisiert werden :

Teilpaket TP 1 : Bereitstellung und Vereinheitlichung der vorhandenen Software ; Detaillierte Einarbeitung in die von beiden Gruppen bislang verfolgten Lösungsansätze ; Umwandlung der in Köln entwickelten Perl-Anwendungen zu TXM-kompatiblen Groovy-Plugins.

Teilpaket TP 2: Erstellung von Werkzeugen zur manuellen Überprüfung und Korrektur der automatisch annotierten Korpusfassungen (vgl. die Arbeitspakete AP 3 bis AP 5 weiter oben). Die Umsetzung dieses Teilpakets ist mit hohem Programmier- und Entwicklungsaufwand verbunden, da sie eine weitreichende Umgestaltung der der Anwendungsplattform TXM zu Grunde liegenden Architektur im Hinblick auf die konsistente Verwaltung mehrerer aufeinander aufbauender Versionen ein und desselben Korpus notwendig macht. Dieser Aufwand erscheint jedoch auf Grund der zeitlich wie qualitativ zu erwartenden Effizienzgewinne hinsichtlich einer kontrollierten und auf Kohärenz bedachten Durchführung der in den Arbeitspaketen AP3 bis AP5 vorgesehenen Annotierungsschritte gerechtfertigt.

Teilpaket TP 3: Erstellung von Extraktions- und Abfragewerkzeugen zur Durchführung von kookkurrenzprofilbasierten Distributionsanalysen.

Dieses Teilpaket betrifft die folgenden Anwendungen und Funktionen :

- Extraktion von Kombinationsmustern auf der Grundlage der – entsprechend der in AP 4 entwickelten Konzeption – durchgeführten syntaktischen Annotierungen

- Extraktion von Kombinationsmustern gemäß sog. « Zonen »²⁰, durch die eine strukturierte Kontextualisierung von Korpusabfragen ermöglicht wird (vgl. Pincemin et al. 2006)
- Erstellung und Verwaltung von Kookkurrenzmatritzen erster und zweiter Ordnung (durch Kotextauszählungen zu einzelnen Pivots ermittelte Lexikogramme sowie auf statistisch gewichteten Lexikogrammauszügen mehrerer Pivots beruhende Ähnlichkeits- und Differenz-Matritzen)

Teilpaket TP 4 : Implementierung multivariater Verfahren

- kritische Bestandsaufnahme zu den von beiden Partnern bereits eingesetzten Analysetechniken (MDS bzw. Korrespondenzanalyse) und hierauf beruhende Auswahl der geeigneten Methoden
- Ausarbeitung von Interpretationsleitlinien bei der Auswertung der auf Grundlage umfangreicher Dimensionsreduzierungen generierten Grafiken

Teilpaket TP 5 : Bereitstellung von textometrischen Werkzeugen zur Erkennung lexikalisch-syntaktischer Musterbildungen

- Implementierung von Berechnungsverfahren zur Bestimmung rekurrenter Kombinationsmuster und kookkurrenzparadigmatischer Ausdrucksinventare

Teilpaket TP 6 : Bereitstellung von textometrischen Werkzeugen zur Erkennung von Dimensionen diachroner Differenzierung

- Implementierung von Werkzeugen zur differentiellen Frequenzberechnung sowie zur automatischen Gruppierung von nach Zeiträumen differenzierten Stichproben anhand multivariater Verfahren.

Arbeitspaket AP 7: Entwicklung einer online zugänglichen Datenbank zur Abfrage und Auswertung der Ergebnisse der linguistischen Analyse

Um eine Gesamtsicht über die Ergebnisse der linguistischen Analyse zu bieten, werden diese in einer relationalen Datenbank (des Typs MySQL) gespeichert, welche auf der Website des Projekts über eine Abfrageschnittstelle zugänglich gemacht wird. Hieraus ergeben sich folgende Teilpakete:

Teilpaket TP 1: Erarbeitung eines Entity-Relationship-Modells gemäß den Analyseparametern der linguistischen Pilotstudie (vgl. 3.2.1.1, AP 1);

Teilpaket TP 2: Einrichtung einer Datenbank des Typs MySQL gemäß dem ausgearbeiteten ER-Modell;

Teilpaket TP 3: Schaffung einer projektinternen Schnittstelle, die die Eingabe der Ergebnisse der linguistischen Analyse ermöglicht;

Teilpaket TP 4: Verknüpfung der Datenbank mit den zur Korpusanalyse eingesetzten Anwendungen (vgl. AP 6 weiter oben);

Teilpaket TP 5: Schaffung einer auf der Website des Projekts zugänglichen Abfrageschnittstelle;

Teilpaket TP 6: Generierung herunterladbarer SQL-Quelldaten mit dem Inhalt der Datenbasis

²⁰ Diese entsprechen tabellenförmig aufbereiteten Extraktionsergebnissen, wie sie etwa im Rahmen der Korpusanwendung TIGERSearch anhand von attributschemadefinierten Abfrageausdrücken ermittelt werden können.

3.2.2 Zeitplan und Aufteilung der Arbeitspakete nach Forschergruppen

3.2.2.1 Programm und Methodik der sprachwissenschaftlichen Untersuchung

Monate	1 - 6		7 - 12		13 - 18		19 - 24		25 - 30		31 - 36	
AP1 : Pilotstudie (Monate : 1-15)	■	■	■	■	■							
AP2 : Erweiterte Umsetzung der entwickelten Methodik (Monate : 16-24)						■	■	■				
AP3 : Umfassende Analyse der Entwicklungen vom klassischen Französisch bis heute (Monate 25-36)									■	■	■	■

Tab. 1 : Zeitplan – Sprachwissenschaftliche Untersuchungen

Verantwortlich : Denis Vigier, Benjamin Fagard, Peter Blumenthal

Für die betreffenden Arbeitspakete vorgesehene Mitarbeiter

Fest angestellte Mitarbeiter :

Lyon : D. Vigier, C. Guillot, **S. Rémi-Giraud**, L. Gardelle, M.-L. Demonet, L. Sarda, B. Fagard, M. Charolles

Köln : P. Blumenthal, S. Diwersy

Nicht fest angestellte Mitarbeiter :

Lyon : 1 Post-Doktorand

Köln : P. Gililov (1 x 0,5 Stellen TVL-13) + 1 Hilfskraft

Vorgesehene Gesamtdauer : 36 Monate

Lieferbare Ergebnisse: nach Zeiträumen differenzierte Übersichtsdarstellungen multivariater Analysen des kombinatorischen Verhaltens französischer Präpositionen ; Übersicht über die Ausdrucksbedürfnisse, die in der Entwicklung der untersuchten Pärpositionalsyntagmen deutlich werden (entsprechend dem weiter oben unter 3.1.1. b) vorgestellten Modell); Sammelband

3.2.2.2 Arbeitsprogramm und Methoden hinsichtlich der Erstellung, Verarbeitung und Veröffentlichung der projektbezogenen Sprachdaten

Monate	1	2	3	4	5-7	8-12	12-24	25	26	27-36
AP1 : Zusammenstellung, Vervollständigung und formatbezogene Normalisierung des zu verwendenden Korpusmaterials	■	■	■	■	■	■				
AP2: Austausch und Abstimmung über Kodierungs- und Annotationskonventionen		■	■	■	■					

Monate	1	2	3	4	5-7	8-12	12-24	25	26	27-36
AP3 : Satz- und Wortsegmentierung, morphosyntaktische Annotation und Lemmatisierung										
AP4 : Syntaktische Korpusannotation										
AP5 : Segmentierung und positionsbezogene Kodierung von satzübergreifenden Texteinheiten										
AP6 : Bereitstellung von Werkzeugen zur textometrischen Korpusauswertung AP7 : Entwicklung einer online zugänglichen Datenbank zur Abfrage und Auswertung der Ergebnisse der linguistischen Analyse										

Tab. 2 : Zeitplan – Erstellung, Verarbeitung und Veröffentlichung der projektbezogenen Sprachdaten

Verantwortlich : Serge Heiden, Sascha Diwersy

Für die betreffenden Arbeitspakete vorgesehene Mitarbeiter

Fest angestellte Mitarbeiter :

Lyon : S. Heiden, B. Pincemin, A. Lavrentiev

Köln : S. Diwersy

Nicht fest angestellte Mitarbeiter :

Lyon: 1 technischer Mitarbeiter für die Anwendungsentwicklung innerhalb der Korpusplattform TXM

Köln : 1 Wissenschaftliche(r) MitarbeiterIn (0,5 Stelle TVL-13) + 1 Hilfskraft

Vorgesehene Gesamtdauer : 36 Monate

Lieferbare Ergebnisse : Arbeitskorpora, textometrische Werkzeuge und Datenbankanwendung

A. Arbeitskorpora

Die zu den jeweiligen Teilkorpora gehörenden französischsprachigen Dokumente sollen den Zeitraum vom 9. bis zum 20. Jahrhundert gleichgewichtet abdecken, Sie zeichnen sich durch die folgenden Bearbeitungsmerkmale aus:

1. mit der Anwendungsplattform TXM kompatible XML-TEI-Formatierung
2. systematisch erfasste Metadaten (gemäß Zeitraum, Textgattung u.a.)
3. Wortformsegmentierung, morphosyntaktische Annotation und Lemmatisierung
4. syntaktische Annotierung

5. positionsbezogene Annotierung von Präpositionalsyntagmen

B. Textometrische Werkzeuge

Im Zuge der projektbedingten Anpassung der Korpusplattform TXM werden folgende Anwendungsmodul neu erstellt oder überarbeitet:

1. Module zum Import von Korpusdokumenten
2. Module zur Segmentierung und automatischen Annotation (PoS, Lemmaformen, syntaktische Konstituenten und Dependenz, Texteinheiten)
3. Module zur manuellen Überprüfung und Korrektur automatischer Annotierungen
4. Module zur Extraktion und Bereitstellung rekurrenter Korpussegmente im Hinblick auf kookkurrenzstatistische Berechnungen
5. Module zur Durchführung kookkurrenzstatistischer Berechnungen
6. Module zur Implementierung multivariater Analysemethoden mit entsprechenden grafischen Visualisierungen
7. Module zur Erkennung der Bildung von Ausdrucksmustern
8. Module zur statistischen Erfassung diachroner Entwicklungen

C. Datenbank mit den Ergebnissen der linguistischen Analyse

1. herunterladbare SQL-Quelldateien
2. auf der Website des Projekts zugängliche Abfrageschnittstelle

Die Arbeitskorpora werden entsprechend dem für die linguistischen Untersuchungen festgelegten Arbeitsprogramm (siehe Abschnitt 3.2.1.1) sukzessive bereitgestellt.

Die Softwarekomponenten werden unter Open-Source-Lizenz frei zur Verfügung gestellt.

3.3 Literaturverzeichnis zum Stand der Forschung, zu den Zielen und dem Arbeitsprogramm

- Amiot, D. & W. De Mulder, (2002), De l'adverbe au préfixe en passant par la préposition : un phénomène de grammaticalisation ?, *Linguisticae Investigationes* XXV/2, 247-273.
- Bekkers, F. (1931), *Etude sur l'emploi que Froissart fait de la préposition*. Dissertation, Groningen.
- Blumenthal, P. (2002), Profil combinatoire des noms : synonymie distinctive et analyse contrastive, *Zeitschrift für französische Sprache und Literatur* 112 : 115-138.
- Blumenthal, P. (2003), Der Begriff der externen und internen Sprachgeschichte in der Romanistik, in *Romanische Sprachgeschichte* 1. Berlin : Teilband, 38-45.
- Blumenthal, P. (2006), *Wortprofil im Französischen* (Beihefte zur ZrP 332), Tübingen.
- Blumenthal, P. (2007), Histoire de la montée en puissance de *début*, in: L. Begioni/C. Muller (s.l.d.): *Problèmes de sémantique et de syntaxe - Hommage à André Rousseau*, Lille : Editions du Conseil Scientifique de l'Université Charles-de-Gaulle - Lille 3, (Travaux et recherches), 445-471.
- Blumenthal, P. (2008), Combinatoire des prépositions : approche quantitative, *Langue Française*, 157, 37-51.
- Blumenthal, P. (2009). "Les noms d'émotion : trois systèmes d'ordre", in: I. Novakova & A. Tutin (eds), *Le Lexique des émotions*, Grenoble : ELLUG 2009 (Langues, Gestes, Paroles), 41-64.

- Blumenthal, P. (2011), Le figement : du XVII^e siècle au français contemporain, in: J.-C. Anscombre/S. Mejri (éds.): *Le figement linguistique: la parole entravée*, Paris: Éditions Champion 2011 (Lexica 18), 283-302.
- Blumenthal, P. (2012, à paraître) : Évolutions de la combinatoire prépositionnelle : le cas de *en*, *Mélanges Olivier Soutet*.
- Blumenthal P., S. Diwersy & J. Mielebacher (2005), Kombinatorische Profile und Profilkontraste: Berechnungsverfahren und Anwendungen, *Zeitschrift für Romanische Philologie* n°121, 1: 49-83.
- Blumenthal, P. & M.-J. Hausmann (2006), *Collocations, corpus, dictionnaires, Langue Française*, 150.
- Blumenthal, P. & S. Mejri (2008) (eds), *Les séquences figées : entre langue et discours*, *Zeitschrift für Französische Sprache und Literatur*, Beihefte 36, Franz Steiner Verlag : Stuttgart.
- Bonami, O. (1999), *Les constructions du verbe : le cas des groupes prépositionnels argumentaux*, Thèse de doctorat, Université Paris 7.
- Bolly, C. (2008), *Les unités phraséologiques : un phénomène linguistique complexe ?* Thèse de doctorat, Université Louvain-la-Neuve.
- Buridant, C. (2000), *Grammaire Nouvelle De L'ancien Français*. Paris : Sedes.
- Butler, C. (2003), *Structure and Function: A Guide to Three Major Structural-Functional Theorie*, Amsterdam, Benjamins.
- Cadiot, P. (1997a), Les paramètres de la notion de préposition incolore, *Faits de langues*, 9, 127-134
- Cadiot, P. (1997b), *Les prépositions abstraites en français*, Paris : Armand Colin.
- Carlier A. & Sarda L. (2010). Le complément de la localisation spatiale : entre argument et adjectif, In F. Neveu, V. Muni Toke, T. Klingler, J. Durand, L. Mondada & S. Prévost (eds.) *CMLF 2010 - 2ème Congrès Mondial de Linguistique Française*.
- Cervoni, J. (1991), *La préposition*, Paris-Louvain-la-Neuve, Duculot.
- Charolles, M. (1997), *L'encadrement du discours : univers, champs, domaines et espaces*, *Cahier de Recherche Linguistique, LANDISCO, URA-CNRS 1035 Université Nancy 2*, n° 6, 1-73.
- Charolles, M. & S. Prévost (2003)(eds.), *Adverbiaux et topiques. Travaux de linguistique* 47.
- Charolles, M. & M.-P. Péry-Woodley (2005)(eds.), *Les adverbiaux cadratifs. Langue française* 148.
- Charolles, M., Vigier, D. (2005), Les adverbiaux en position préverbale : portée cadrative et organisation des discours. *Langue Française*, 148, 9-30.
- Coseriu, E. (1974), *Synchronie, Diachronie und Geschichte. Das Problem des Sprachwandels*, München.
- Croft, W. & A. Cruse (2004), *Cognitive Linguistics*, Cambridge, CUP.
- Dannell, K. J., Nothing but phrases. About the distribution of idioms and stock phrases, in Edlund, L.E. & Persson, G. (éds). *Language: the time machine*, Umeå : Umeå University
- Dendale, P. & W. De Mulder, (1998), *Contre et sur : du spatial au métaphorique ou inversement ?*, *Verbum* XX/4, 405-434.
- De Mulder, W. & D. Stosic (2009) *Présentation*, in De Mulder, W. & Stosic, D. (2009) (eds) *Approches récentes de la préposition, Langages*, 173, 3-13.
- De Mulder, W. & A. Vanderheyden (2002), The evolution of French 'sur' : steps toward a diachronic prototype semantics of spatial prepositions, In Hubert Cuyckens (ed.), *Perspectives on prepositions*. Tübingen : Niemeyer, 177-193.
- Denis, P., Sagot B. (2009), « Coupling an annotated corpus and a morphosyntactic lexicon for state-of-the-art POS tagging with less human effort », in *Proceedings of The Pacific Asia Conference on Language, Information and Computation (PACLIC 23)*, Hong Kong, China.

- Dufresne, M., F. Dupuis & M. Tremblay (2003), Les prépositions dans l'histoire du français : transitivité, grammaticalisation et lexicalisation, *Verbum* XXV/4, 551-564.
- Everaert, M. et al.(eds). 1995. *Idioms*. Hillsdale, N.J.: Erlbaum.
- Fagard, B. (2004), *Après : de l'espace au temps, la sémantique en diachronie*. In A. Soares da Silva, A. Torres & M. Gonçalves (eds.), *Linguagem, cultura e cognição*, Braga : Almedina, 231-246.
- Fagard, B. (2006), La grammaticalisation en question – du latin aux langues romanes modernes, *Modèles linguistiques* XXVII-1/53, 91-110.
- Fagard, B. (2007), De *puis* à *depuis* : préfixation et évolution sémantique. In B. Combettes & C. Marchello-Nizia (eds.), *Etudes sur le changement linguistique du français*, 129-143.
- Fagard, B. (2008), 'Côté' dégrammaticalisation – le cas des prépositions. In B. Fagard, S. Prévost, B. Combettes & O. Bertrand (eds.), *Evolutions en français*, Berne : Peter Lang.
- Fagard, B. 2010. *Espace et grammaticalisation - L'évolution sémantique des prépositions dans les langues romanes*. Sarrebruck : Editions Universitaires Européennes.
- Fagard, B. & W. De Mulder (2007), La formation des prépositions complexes : grammaticalisation ou lexicalisation ?, *Langue française* 156, 9-29.
- Fagard, B. & L. Sarda (2009), Etude diachronique de la préposition *dans*, in J. François, É. Gilbert, C. Guimier & M. Krause (eds.), *Autour de la préposition*, Caen : Presses Universitaires de Caen, 225-236.
- Fournier, N. (1998), *Grammaire du français classique*, Paris : Belin
- Franckel J.J. & Paillard D. (2006), De la couleur des prépositions dans leurs emplois fonctionnels, *Modèles Linguistiques*, XXVII, 2, 54, 51-56.
- Franckel J.J. & Paillard D. (2008), *Grammaire des prépositions*, Paris : Ophrys.
- Gougenheim, G. (1974), *Grammaire de la langue française du 16e siècle*. Paris : Picard.
- Gross, G. (1996), *Les expressions figées en français. Noms composés et autres locutions*, Paris : Ophrys.
- Grossmann, F. & A.Tutin (2003) (eds), *Les collocations. Analyse et traitement*, Amsterdam : De Werelt.
- Grunig, B.-N. (1997), La locution comme défi aux théories linguistiques : une solution d'ordre mémoriel?, in Martins-Baltar, M. (éd.), *La locution entre langue et usages*, Paris Fontenay-St Cloud : ENS, 225-240.
- Heine B., U. Claudi & F. Hünemeyer (1991), *Grammaticalization : A Conceptual Framework*. Chicago : University of Chicago Press.
- Heine, B. & T. Kuteva, (2002), *World Lexicon of Grammaticalization*. Cambridge, Cambridge University Press.
- Hopper P. & E. Traugott (éds) (1993), *Grammaticalization*. Cambridge University Press.
- Hoey, M. (2005), *Lexical Priming*, London, Routledge
- Homles, M. (2008) « Domestic strife in early modern Europe: images and texts in a virtual anthology ». (Conference presentation) with Claire Carlin; Digital Humanities 2008.
- Homles, M. (2010) « Using the Universal Similarity Metric to Map Correspondences between Witnesses ». (Conference presentation) Digital Humanities 2010. Conference, Kings College London..
- Hudson, J. (1998), *Perspectives on fixedness: applied and theoretical*, *Lund Studies in English*, 94. Lund: Lund University Press.
- Hunston, S. & G. Francis, (2000), *Pattern Grammar*, Amsterdam: Benjamins.
- Kilroe, P. (1989), *The grammaticalization of French à*. Doctoral dissertation, University of Texas : Austin.
- Kilroe, P. (1994), The grammaticalization of French à. In William Pagliuca (ed.), *Perspectives on Grammaticalization*, Amsterdam / Philadelphia : J. Benjamins, 49-61.
- Lamiroy, B., (2010), *Les expressions verbales figées de la francophonie. Les variétés de Belgique, de France, du Québec et de Suisse*. Paris :Ophrys.

- Lamiroy, B. & J. Klein (2010), Le problème central du figement est le semi-figement, in *Linx*, 53 (2005), 135–154.
- Lang, J. (1991), *Die französischen Präpositionen. Funktion und Bedeutung*, Heidelberg.
- Lavieu, B. (2006), Distinction entre groupes prépositionnels régis et non régis. *Modèles linguistiques* 53, tome XXVII-1 : 131-144.
- Lebel, M-É. (2003), *Le morphème par en ancien français*. Thèse en linguistique, Université du Québec à Montréal
- Lecolle, M. & S. Leroy (2006) (eds), *Changements linguistiques : figement, lexicalisation, grammaticalisation, Cahiers de praxématique*, 46 .
- Le Draoulec, A. & Vigier, D. (2009). Dans la nuit je rentre chez moi. *Interprétations temporelle et spatiale de quelques constructions dans + SNtps*. Actes des cinquièmes Rencontres de Sémantique et Pragmatique - « Espace, Temps dans les langues », *Revue de sémantique et de pragmatique*.
- Leeman, D. (1998), *Les circonstants en question(s)*, Paris, Kimé.
- Leeman, D. (1999), L'unité lexicale dans la perspective harrissienne, *LINX* 40 : 117-136.
- Leeman, D. (2006), la préposition française : caractérisation syntaxique de la catégorie, in Leeman, D. (2006) (éd), *la préposition en français, I*, *Modèles linguistiques* 54, tome XXVII, 7-17.
- Leeman, D. (2008), Prépositions du français, état des lieux, in Leeman, D. (2008) (éd), *Enigmatiques prépositions, Langue Française*, 157, 5-19.
- Le Pesant, D. (2006), Classification à partir des propriétés syntaxiques, *Modèles linguistiques* 53, tome XXVII-1 : 51-74.
- Marchello-Nizia, C. (1997 [1979]), *Histoire de la langue française aux 14^{ème} et 15^{ème} siècles*. Paris : Bordas.
- Marchello-Nizia, C. (1999), *Le Français en diachronie*, Paris : Ophrys.
- Marchello-Nizia, C. (2002), Prépositions françaises en diachronie : une catégorie en question, *Linguisticae Investigationes* XXV/2, 205-221.
- Martineau, F. (2000), Les prépositions à et de de l'ancien français au français classique. In Claude Buridant (dir.), *Actes du IXe colloque international sur le moyen français*. Strasbourg : Presses Universitaires de Strasbourg, 165-183.
- Marchello-Nizia, C. (2006), *Grammaticalisation et changement linguistique*. Bruxelles : Duculot / De Boeck
- Mejri, S. (2000), Figement lexical et renouvellement du lexique : quand le processus détermine la dynamique du système, *Le français moderne* LXVIII (1), 39-62.
- Melis, L. (2003), *La préposition en français*, Paris : Ophrys.
- Mel'čuk, I. (2003), Collocations : définition, rôle et utilité, In F.Grossman et A.Tutin (éds), 23-31.
- Michael, F. (1904), *Über den Gebrauch der Präpositionen mit Ausschluss von a, de, en, im altfranzösischen Rolandsliede*. Dissertation, Greifswald.
- Moon, R. (1998), *Fixed Expressions and Idioms in English, a Corpus-Based Approach*, Oxford : Clarendon
- Noailly, M. (2010), L'adjectif, quand il est invariable, sort-il pour autant de sa catégorie ?, in J. Goes & E. Moliné (eds), *L'adjectif hors de sa catégorie*, Actes du sixième colloque franco-roumain, Université d'Artois, 23-25 mai 2007, Arras.
- Raithel, G. (1875), *Über den Gebrauch und die begriffliche Entwicklung der altfranzösischen Präpositionen od, par, en*. Dissertation, Göttingen.
- Sagot B. (2010), « The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French », in *Proceedings of the 7th Language Resources and Evaluation Conference (LREC'10)*, Vallette, Malta.
- Sarda L. & Charolles M. (eds) (soumis) Les adverbiaux prépositionnels : position, fonction et portée. Numéro thématique, *Travaux de Linguistique*.
- Senellart, J. (1998), Reconnaissance automatique des entrées du lexique-grammaire des phrases figées, in Lamiroy, B. (éd), *Le lexique-grammaire. Travaux de Linguistique*, 37, numéro spécial. 109-127.

- Shears, F. (1922), *Recherches sur les prépositions dans la prose du moyen français (XIVe et XVe siècles)*. Paris : Champion.
- Shyldkrot, H. B.-Z. & S. Kemmer (1995), La grammaticalisation des prépositions : concurrence et substitutions, *Revue Romane* 30, 205-226.
- Stein, A. (2003), « Étiquetage morphologique et Lemmatisation de textes d'Ancien Français. » in: *Ancien et Moyen Français sur le Web: enjeux méthodologiques et Analyse du Discours*, P. Kunstmann & al. (eds), Ottawa: David Éditions, 273–284.
- Stein, A. and al. (eds) (2006), *Nouveau Corpus d'Amsterdam. Corpus informatique de textes littéraires d'Ancien Français (1150-1350)*, texte établi par A. Dees (Amsterdam 1987), remanié par A. Stein, P. Kunstmann & M.-D. Glessgen. Stuttgart: Institut für Linguistik/Romanistik.
- Svensson, M.H. (2004), *Critères de figement. L'identification des expressions figées en français contemporain*. Umeå : Umeå University.
- Thompson, S. (1985), Grammar and written discourse: Initial vs. final purpose clause in English, *Text* 5(1-2), 55-84.
- Traugott E.C. & König E. (1991), The Semantics-Pragmatics of Grammaticalization Revisited, in E.C. Traugott & B. Heine (eds), *Approaches to Grammaticalization*, vol. 1, Amsterdam/Philadelphia, John Benjamins, 189- 218.
- Vaguer, C. (2006), Bibliographie générale des prépositions du français (I et II). *Modèles linguistiques* 54, tome XXVII-II : 171-203.
- Vaguer, C. (2007), Bibliographie générale des prépositions du français (III). *Modèles linguistiques* 55, tome XVIII-I.
- Vaguer, C. (2008), Classement syntaxique des prépositions simples du français, *Langue Française*, 157, 20-36.
- Vigier, D. (2003), Les syntagmes prépositionnels en « en N » détachés en tête de phrase référant à des domaines d'activité, *Linguisticae Investigationes*, 26 (1), 97 - 122, Amsterdam, John Benjamins B.V.
- Vigier, D. (2004), *Les groupes prépositionnels en « En N » : de la phrase au discours*, Thèse de 3^e cycle, Université Paris 3.
- Vigier, D. (2005). Les adverbiaux praxéologiques détachés en position initiale et leur portée. *Verbum*, XXVII, 2005, 3, 293-312.
- Vigier, D. (2008a). Les adverbiaux de localisation spéculaire : détachement et portée. In D. Apothéloz, B. Combettes, F. Neveu (eds) *Les linguistiques du détachement. Actes du colloque international de Nancy (7-9 juin 2006)*. Berne : P. Lang, 567-580.
- Vigier, D. (2008b), Contribution à une étude des constructions antéposées du type : *En homme intelligent et humain, il partagea tout de suite l'inquiétude de Marcel* (J. Verne), *Discours*, 2.
- Vigier D. (2008c). La gestion des cadres de discours dans une tâche rédactionnelle en FLE. In H. Hilton (ed). *Acquisition et didactique 1. Actes de l'atelier didactique, colloque international AFLS 2005*. Chambéry: Presses Universitaires de Chambéry, 113-129.
- Vigier, D. (à par., 2012a), « Comportements, déguisements, rôles, ... De quelques emplois intraprédicatifs de en », *Linguisticae Investigationes*, Amsterdam, Philadelphia : John Benjamins Publishing Co.
- Vigier, D. (à par., 2012b), « En attendant : un cas de pragmatization », in *Les adverbiaux prépositionnels : position, fonction et portée*, Charolles, M. & Sarda, L. (éds.), *Travaux de linguistique*.
- Vigier, D. (ed.) (en prép.), *La préposition « en »*, *Langue Française*, paris : Hachette
- Vigier, D., Terran E. (eds) (2005). *Les adverbiaux cadratifs et l'organisation des textes*. *Verbum*, XXVII, 3.
- Virtanen, T. (1992a.), *Discourse functions of adverbial placement in English*. Åbo: Åbo Akademi University Press.
- Virtanen, T. (1992b), Temporal Adverbials in Text Structuring: On Temporal Text Strategy. In A.-C. Lindberg, N. Enkvist & K. Wikberg (eds.), *Nordic Research on Text and Discourse*. Åbo: Åbo Academic Press, 85-197.

Wray, A. (2002), *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

4. Beantragte Mittel/Module

4. 1 Basismodul

4.1.1 Personalkosten

4.1.1.1 Nach TV-L vergütetes Personal

a) Postdoktorand

1 halbe Stelle Postdoktorand/TVL-E 13 Stufe 4	36 Monate	89.550 €
---	-----------	----------

b) Doktorand

1 halbe Stelle Doktorand/TVL-E 13 Stufe 1	36 Monate	84.150 €
---	-----------	----------

4.1.1.2 Hilfskräfte

zwei studentische Hilfskräfte/30 Std./Monat	36 Monate	52.000 €
---	-----------	----------

Die studentischen Hilfskräfte sollen bei der Zusammenstellung und Aufbereitung der Korpora (Lemmatisierung, PoS-Tagging, Parsing), beim Training von Software, bei der Literaturrecherche und -beschaffung, bei der Einrichtung der Datenbank und bei der verlegerischen Vorbereitung der Veröffentlichungen helfen.

Personalkosten Gesamt: 225.700 €

4.1.2 Sachmittel

4.1.2.1 Geräte bis 10.000,- EUR, Software und Verbrauchsmaterial

4.1.2.2 Reisemittel

3 Reisen nach Lyon (3 Personen) à 750 € .p.P./Reise	6.750€
---	--------

3 Reisen nach Lyon (1 Person) à 750 € .p.P./Reise	2.250€
---	--------

3 x Teilnahme an Tagung Ausland (4 Personen) à 750 € p.P./Reise	9.000€
---	--------

4 x Teilnahme an Tagung Inland (4 Personen) à 350 € p.P./Reise	5.600€
--	--------

Reisekosten Gesamt 23.600 €

4.1.2.3 Mittel für wissenschaftliche Gäste

4.1.2.4 Mittel für Versuchstiere

4.1.2.5 Sonstige Mittel

4.1.2.6 Projektinterne Publikationsmittel

Sammelband bei John Benjamins, Amsterdam/Philadelphia, PA	3.000€
---	--------

(pauschal)

4.2 Modul Eigene Stelle

4.3 Modul Vertretung

4.4 Modul Mercator-Fellow

4.5 Modul Projektspezifische Workshops

Durchführung eines eintägigen Workshops in Köln (2015)

16.000€

4.6 Modul Öffentlichkeitsarbeit

5. Überblick über die beantragten Mittel

	FR		ALL
Dépenses de personnel	192.358	Basismodul: Personalkosten	225.700
Equipements		Basismodul: Wissenschaftliche Geräte	
Frais de missions	60.040	Basismodul: Reisekosten	23.600
Autres dépenses	12.000	Basismodul: Sonstige Kosten	
-		Basismodul: Publikationskosten	3.000
Prestations de service	25.000	-	
Frais de gestion	11.576	-	
-		Modul Projektspezifische Workshops	16.000
		Modul Öffentlichkeitsarbeit	
Total France	€ 300.974	Gesamtsumme Deutschland	€ 268.300
TOTAL	€ 569274		

6. Voraussetzungen für die Durchführung des Vorhabens

Keine weiteren Voraussetzungen sind außer der Grundausstattung gegeben.

7. Erklärungen

Ein Antrag auf Finanzierung dieses Vorhabens wurde bei keiner anderen Stelle eingereicht. Wenn wir einen solchen Antrag stellen, werden wir die Deutsche Forschungsgemeinschaft und die Agence Nationale de la Recherche unverzüglich benachrichtigen.

Der Vertrauensdozent der DFG an der Universität zu Köln wurde von der Antragstellung unterrichtet.

8. Unterschriften



(Denis Vigier)



(Peter Blumenthal)

II. ANHANG

II.1. Lebensläufe aller Projektbeteiligten

BLUMENTHAL Peter

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: BLUMENTHAL
Vorname: Peter
Alter: 66 Pensionierung am 31.7.2013)
Gegenwärtige Stellung: Universitätsprofessor

Akademische Ausbildung und beruflicher Werdegang :

1965-69: Studium der Romanistik und allgemeinen Sprachwissenschaft an den Universitäten Bonn und Köln
1967/68: Studium in Aix-en-Provence
1969: Staatsexamen
1970: Promotion in Romanistik: *Die Entwicklung der romanischen Labialkonsonanten*, Rheinische Friedrich-Wilhelms-Universität Bonn
1970-72: assistant associé an der Universität Toulouse-Le Mirail
1972-78: wissenschaftlicher Assistent an der Universität zu Köln
1978: Habilitation: *La syntaxe du message. Application au français moderne*, Universität zu Köln
1983-88: Univ.-Prof., Friedrich-Alexander Universität Erlangen-Nürnberg
1988-98: Univ.-Prof., Universität Stuttgart
1990: Gastprofessor in Aarhus (Dänemark)
1991: Dekan der philosophischen Fakultät, Universität Stuttgart
1992: Gastprofessor an der Universität Laval (Québec)
1992-2000: Sachverständiger der Deutschen Forschungsgemeinschaft für romanische Sprachwissenschaft
1998-: Rufannahme an der Universität zu Köln
2001: Februar-März, professeur associé in Paris 7
2003: nomination en qualité d'expert étranger de l'Institut de linguistique française
2007: März, professeur invité in Paris 13
2007: nomination en qualité d'expert des sciences humaines et sociales (DGRI-A6) par le Ministère de l'enseignement supérieur et de la recherche.
2010: Directeur de recherche associé au CNRS (3 mois)

B/ Weitere berufliche Erfahrungen

- Herausgeber einer linguistischen Fachzeitschrift (*Zeitschrift für französische Sprache und Literatur*) sowie mehrerer Sammlungen;
- Leiter des Fremdspracheninstituts der Universität Stuttgart (in den 90er Jahren)
- Président du comité d'évaluation de l'ATILF, Nancy (janvier 2012)

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

- 1) *Wortprofil im Französischen* (Beihefte zur ZrP 332), Tübingen 2006.
- 2) "Histoire de la montée en puissance de *début*", in: L. Begioni/C. Muller (s.l.d.): *Problèmes de sémantique et de syntaxe - Hommage à André Rousseau*, Lille: Editions du Conseil Scientifique de l'Université Charles-de-Gaulle - Lille 3 2007 (Travaux et recherches), 445-471.
- 3) "Combinatoire des prépositions: approche quantitative", in: *Langue française* 157 (2008), 37-51.
- 4) "La sémantique cognitive face à l'idiomaticité", in: J. François (s.l.d.): *Grandes voies et chemins de traverse de la sémantique cognitive*, Leuven: Peeters 2010 (Mémoires de la Société de Linguistique de Paris, Nouvelle série, Tome XVIII), 147-

162.<http://www.linguistiquefrancaise.org/articles/cmlf/pdf/2008/01/cmlf08338.pdf>

- 5) "Le figement : du XVII^e siècle au français contemporain", in: J.-C. Anscombre/S. Mejri (éds.): *Le figement linguistique: la parole entravée*, Paris: Éditions Champion 2011 (Lexica 18), 283-302.

VIGIER Denis

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: VIGIER

Vorname: Denis

Alter: 50

Gegenwärtige Stellung: Maître de conférences en linguistique française, Université Lyon 2 Lumière.

Akademische Ausbildung und beruflicher Werdegang :

1989: Agrégation de Lettres

1989-1999: Professeur dans le secondaire

1999-2007: Professeur Agrégé (PRAG) à l'Université Rennes 2 (Haute Bretagne)

2004: Thèse de Sciences du Langage, Université Paris 3 « *Les groupes prépositionnels en « en N »: de la phrase au discours* »

2007-: Maître de conférences en linguistique française, Université Lyon 2 Lumière

B/ Weitere berufliche Erfahrungen

Co-rédacteur en chef de la revue discours (<http://discours.revues.org/>) depuis 2009

Membre de comités scientifiques pour divers colloques internationaux

Membre expert nommé par l'Aeres pour l'évaluation de l'EA Crisco (octobre 2010)

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

Vigier, D. (à par., 2012), « Comportements, déguisements, rôles, ... De quelques emplois intraprédicatifs de *en* », *Linguisticae Investigationes*, Amsterdam, Philadelphia : John Benjamins Publishing Co.

Vigier, D. (à par., 2012), « *En attendant* : un cas de pragmatization », in *Les adverbiaux prépositionnels : position, fonction et portée*, Charolles, M. & Sarda, L. (éds.), *Travaux de linguistique*.

Vigier, D. (2008). Les adverbiaux de localisation spéculaire : détachement et portée. In D. Apothéloz, B. Combettes, F. Neveu (eds) *Les linguistiques du détachement. Actes du colloque international de Nancy (7-9 juin 2006)*. Berne : P. Lang, 567-580.

Vigier, D. (2008). Contribution à une étude des constructions antéposées du type : *En homme intelligent et humain, il partagea tout de suite l'inquiétude de Marcel*. (J. Verne) ». *Discours*, 2. (<http://discours.revues.org/index863.html>)

Vigier D. (2008). La gestion des cadres de discours dans une tâche rédactionnelle en FLE. In H. Hilton (ed). *Acquisition et didactique 1. Actes de l'atelier didactique, colloque international AFLS 2005*. Chambéry: Presses Universitaires de Chambéry, 113-129.

- Vigier, D. (2005). Les adverbiaux praxéologiques détachés en position initiale et leur portée. In Vigier, D., Terran É. (eds), *Les adverbiaux cadratifs et l'organisation des textes*, *Verbum*, XXVII, 2005, 3, 293-312)
- Charolles, M., Vigier, D. (2005). Les adverbiaux en position préverbale : portée cadrative et organisation des discours. *Langue Française*, 148, 9-30.

CHAROLLES Michel

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: CHAROLLES

Vorname: Michel

Alter: 64

Gegenwärtige Stellung: PRCE, université de Paris III

Akademische Ausbildung und beruflicher Werdegang:

Doctorat d'Etat en linguistique à l'Université de Besançon en 1982. Assistant à l'Université de Rennes II de 1983 à 1987. Professeur à l'Université de Nancy II (1987-1997), puis à l'université de Paris III.

B/ Weitere berufliche Erfahrungen

M. Charolles a participé à de nombreux projets sur la référence, les adverbiaux et la structuration des discours. Il a notamment coordonné un projet PIR Cognisciences (1991-1993) sur l'anaphore et un projet ANR Blanc sur les adverbiaux (2006-2009). A Paris III, il a présidé la Commission de Spécialistes de 7ème section (2007-2008) et dirigé l'Ecole Doctorale 268 « Langues et Langage » (2007-2008). Depuis 2008, il est directeur de l'UMR 8094 LATTICE (« Langues, Textes, Traitements Informatiques et Cognition ») dont l'établissement support est l'ENS Paris, avec Paris III comme rattachement secondaire.

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

CHAROLLES M. & PREVOST S., eds, 2003. Adverbiaux et topiques. *Travaux de linguistique*, 47.

CHAROLLES M. & PERY-WOODLEY M-P., eds., 2005, "Les adverbiaux cadratifs", *Langue Française*. 148

CHAROLLES M., 2005, "Framing Adverbials and their Role in Discourse Cohesion: From Connection to Forward Labelling", in M.Aurnague, M.Bras, A Le Draoulec & L.Vieu, *SEM-05 Proceedings*, Biarritz, 13-30.

CHAROLLES M., 2006, "Un jour (one day) in narratives", in I.Korzen & L.Lundquist (eds.), *Comparing Anaphors. Between Sentences, Texts and Languages*, Copenhagen, *Copenhagen Studies in Language*, 34, Copenhagen, Samfundslitteratur Press: 11-26.

CHAROLLES M. & LAMIROY B., 2007, «Du lexique à la grammaire : seulement, simplement, uniquement», *Cahiers de lexicologie*, 90 : 93-116.

DEMONET Marie-Luce

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: DEMONET

Vorname: Marie-Luce

Alter: 59

Gegenwärtige Stellung: Professeur de littérature française, Université de Tours

Akademische Ausbildung und beruflicher Werdegang:

Spécialiste des rapports entre littérature, langues et théories sémiotiques à la Renaissance (*Les Voix du signe*, 1992), M.-L. Demonet a enseigné dans les universités d'Aix-en-Provence, Clermont-Ferrand, Poitiers et Tours (2001). Elle a publié plusieurs études sur Rabelais et une transcription de ses oeuvres en ligne (1995), parallèlement à ses travaux sur Montaigne (*Les Essais*, 1985, *Montaigne et le Nouveau Monde*, ed., 1992, *Montaigne et la question de l'homme*, éd. 1999, *A plaisir. Sémiotique et scepticisme chez Montaigne*, 2002, *L'écriture du scepticisme chez Montaigne*, co-éd. 2003). Elle étudie les transformations du vocabulaire intellectuel, juridique et social et les questions de théorie et d'esthétique littéraires : sur le statut de la fiction et sur la conception des signes à la Renaissance. Elle a été directrice du Centre d'Etudes Supérieures de la Renaissance de Tours (2003-2007), et elle est responsable du programme des « Bibliothèques Virtuelles Humanistes » (BVH) depuis 2003, pour la publication en ligne d'ouvrages du XVIe siècle et leur traitement par de nouveaux outils d'encodage et d'indexation (<http://www.bvh.univ-tours.fr>). En décembre 2010, les BVH ont obtenu une bourse Google pour les traitements linguistiques des bases en français de la Renaissance (régularisation et requêtes). M.-L. Demonet est membre de l'Institut Universitaire de France depuis 2005 (statut renouvelé en 2010).

B/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

1. 1995, *Electro-chroniques de Rabelais*, CD-ROM. Sélection des textes romanesques et parabelaisiens, et établissement des liens en hypertexte (avec E. Brunet), Paris, éditions Les Temps qui Courent. Version en ligne (par Etienne Brunet) : <http://134.59.31.3/Rabelais.html>. En cours de migration sur le site des BVH.
2. 2004, « La Renaissance numérique: les Bibliothèques Virtuelles Humanistes », Euroconférence *Digital technology and philological disciplines*, éd. J.-L. Lebrave and A. Bozzi, Castelvechchio Pascoli, *Linguistica computazionale*, XX-XXI, 2004, p. 175-185.
3. 2008, Nicole Dufournaud, avec la collaboration de Marie-Luce Demonet et Toshinori Uetani, *Manuel d'encodage XML-TEI, Renaissance et temps modernes*, sur [bvh.univ-tours.fr](http://www.bvh.univ-tours.fr). (révision et version 2 en juillet 2009) <http://www.bvh.univ-tours.fr/XML-TEI/index.asp>
4. 2008, Marie-Hélène Lay, Marie-Luce Demonet, « Some elements of reflection about sustainability and shareability of the Humanist Virtual Library (Bibliothèques Virtuelles Humanistes, BVH): a synthesis and experiment feedback », communication acceptée à l'atelier « Sustainability of Language Resources and Tools for Natural Language Processing », *LREC Conference, 6th Language Resources and Evaluation Conference*, Marrakech. <http://www.lrec-conf.org/proceedings/lrec2008/>
5. 2009, communication au congrès des « Digital Humanities », Université du Maryland, « The 'Bibliothèques Virtuelles Humanistes' (Virtual Humanistic Libraries in Tours): a Collection, or a Corpus ? », http://mith2.umd.edu/dh09/index.html%3Fpage_id=99.html.

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: DIWERSY

Vorname: Sascha

Alter: 41

Gegenwärtige Stellung: wissenschaftlicher Assistent / wissenschaftlicher Lehrbeauftragter

Akademische Ausbildung und beruflicher Werdegang:

1991-93: Studium der Angewandten Linguistik/Übersetzungswissenschaft (Französisch, Russisch), Universität des Saarlandes, Saarbrücken

1994/95: assistant-enseignant d'allemand au Lycée Kléber à Strasbourg

1996-99: Studium der Angewandten Linguistik/Übersetzungswissenschaft (Französisch, Russisch), Universität des Saarlandes, Saarbrücken

2000: Magister Artium, Universität des Saarlandes, Saarbrücken

2002-: assistant de recherches à l'Université de Cologne

2008 : Promotion in Romanistik: *Kookurrenz, Kontrast, Profil: Korpusinduzierte Studien zur lexikalisch-syntaktischen Kombinatorik französischer Substantive (mit ergänzenden Betrachtungen zum Deutschen)* Universität zu Köln

2009-: enseignant-chercheur au Centre de recherche interdisciplinaire sur la France et la Francophonie (CIFRA) de l'Université de Cologne

B/ Weitere berufliche Erfahrungen

1999-2000 : Lehrbeauftragter, Fremdspracheninstitut Saarbrücken

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

Diwersy S. (à par., 2012) "La francophonie multivariée. Ou : comment mesurer les français en Afrique ?", in: Blumenthal, P., Pfänder, S. (eds.): *Convergences, divergences et question de la norme en Afrique francophone* (= Le français en Afrique 27

Diwersy S., Evert S., Neumann S. (à par., 2012) "A semi-supervised multivariate approach to the study of language variation", in: Szmrecsanyi, B., Waelchli, B. (eds.) (2012): *Aggregating dialectology and typology: linguistic variation in text and speech, within and across languages*, Berlin e.a.: de Gruyter.

Diwersy S., Eufe R. (2011) "Die karolingische correctio und der lateinische Wortschatz – Kommunikationsverben in Heiligenviten als Fallbeispiel", in: Dessì Schmid, S., Hafner, J., Heinemann, S. (eds.): *Koineisierung und Standardisierung in der Romania*, Heidelberg: Winter, p. 111-131.

Diwersy S. (2007) *Kookurrenz, Kontrast, Profil: Korpusinduzierte Studien zur lexikalisch-syntaktischen Kombinatorik französischer Substantive (mit ergänzenden Betrachtungen zum Deutschen)*, thèse de doctorat non publiée, Université de Cologne, 2007).

Diwersy S., Blumenthal P., Mielebacher J. (2005) "Kombinatorische Profile und Profilkontraste: Berechnungsverfahren und Anwendungen", in : *Zeitschrift für Romanische Philologie* 121, 1 49-83.

FAGARD Benjamin

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: FAGARD

Vorname: Benjamin

Alter: 33

Gegenwärtige Stellung: Chargé de Recherche au CNRS (laboratoire Lattice, CNRS & ENS)

Akademische Ausbildung und beruflicher Werdegang:

1996-2000: Etudes de lettres et linguistique des langues classiques (latin et grec) aux Universités de Paris 4 et Paris 10 (1996-2000).

2000: Agrégation de grammaire

2001: DEA de linguistique théorique à l'Université Paris 7

2006: Thèse en linguistique théorique à l'Université Paris 7 et Université Rome 3 (cotutelle)

B/ Weitere berufliche Erfahrungen

Enseignement en tant que titulaire (moniteur-allocataire ou ATER, période 2001-2007):

- Université Paris 3 (ancien français)

- Université Paris 7 (linguistique théorique, syntaxe, morphologie, diachronie, ancien français, littérature médiévale)

- Université de Valenciennes (ancien français, littérature médiévale, théorie littéraire, critique littéraire)

Enseignement en tant que chargé de cours (période 2007-2012) :

- Université de Rennes 2 (grammaire du 16ème et du 18ème pour l'agrégation)

- Université de Cayenne (ancien français pour le CAPES)

- Université Catholique de Louvain (la grammaticalisation, séminaire de Master de Liesbeth Degand)

- Université de Neuchâtel (grammaticalisation et langues romanes, séminaire de Master de Gilles Corminboeuf)

- Université Paris 3 (écrit et oral en linguistique française, séminaire de thèse de Florence Lefeuvre)

- Université Paris 7 (grammaticalisation et langues romanes, séminaire de thèse de Sophie Prévost)

- CIEP (linguistique et typologie)

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

Fagard, B. & de Mulder, W. 2010. *Devant : Evolution sémantique d'une préposition en français*, CMLF2010 Proceedings. DOI: 10.1051/cmlf/2010185

Fagard, B. 2010. *Espace et grammaticalisation - L'évolution sémantique des prépositions dans les langues romanes*. Sarrebruck : Editions Universitaires Européennes.

Fagard B. 2009. *Vers, du français classique au français moderne ou contemporain : l'aboutissement d'une grammaticalisation ?*, *Revue québécoise de linguistique* 33/1, 5-29.

Prévost, S. & Fagard, B., 2007. Grammaticalisation et lexicalisation : la formation d'expressions complexes, *Langue française* 156, 3-8.

Fagard, B., 2006. La grammaticalisation en question – du latin aux langues romanes modernes, *Modèles linguistiques* XXVII-1/53, 91-110.

GARDELLE Laure

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: Gardelle

Vorname: Laure

Alter: 31

Gegenwärtige Stellung: Maître de conférences en linguistique anglaise, ENS Lyon ;
membre du laboratoire ICAR.

Akademische Ausbildung und beruflicher Werdegang:

- 1996-1998 DEUG LLCE anglais, Université Lyon III, mention très bien
1998-1999 Licence LLCE anglais, Lyon III, mention très bien
1999-2000 - Maîtrise LLCE anglais, Lyon III, mention très bien
- Lectrice en Angleterre (University of Leeds)
2000-2001 - Agrégation d'anglais option linguistique (reçue 3^{ème})
- Capes d'anglais (reçue 5^{ème})
2001-2002 - DEA de linguistique, Lyon III : « Les animaux et le genre en anglais moderne : le système des pronoms personnels ». Mention très bien.
- Stage d'agrégation (lycée Jean-Paul Sartre, Bron)
2002-2006 Thèse de linguistique, sous la direction du professeur P. Cotte : « Le genre en anglais moderne (XVI^e siècle à nos jours) : le système des pronoms ». Soutenue en l'université Paris IV-Sorbonne le 1^{er} décembre 2006. Mention très honorable avec félicitations du jury à l'unanimité. (1200 pages, dont 750 de rédaction ; corpus diachronique de 5200 occurrences)
2002-2007 Professeur agrégé d'anglais (PRAG), Université Lyon III
2007-2009 Maître de conférences en linguistique anglaise, Université Lyon III
2009- Maître de conférences en linguistique anglaise, ENS Lyon : linguistique et phonologie de l'anglais (licence, master 1, master 2, agrégation)

B/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

Gardelle, L. (2011) "Whoop her up, hit it, go it alone: the role of the personal pronoun in the fossilization process". Yearbook of Phraseology 2011, pp. 163-178.

GUILLOT Céline

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: GUILLOT

Vorname: Céline

Alter: 41

Gegenwärtige Stellung: Maître de conférences à l'ENS de Lyon (7e section), membre junior de l'Institut universitaire de France

Akademische Ausbildung und beruflicher Werdegang:

- septembre 2005 – février 2006 : détachement au CNRS sur un poste de CR2 dans l'UMR

ICAR

- septembre 2004 – août 2005 : congé parental

- septembre 1999 – août 2004 / ATER à l'ENS Lettres et sciences humaines, UMR ICAR

- 1993 – 1994 : Scolarité à l'École nationale du Patrimoine

- 1989 – 1992 : Scolarité à l'École nationale des Chartes

B/ Weitere berufliche Erfahrungen

- 1998-1999 : Conseillère pour les archives à la Direction Régionale des Affaires Culturelles de Rhône-Alpes

- 1994-1998 : Conservateur adjoint aux Archives Départementales de la Loire

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

(2004) « *Ceste parole et ceste aventure dans la Queste del Saint Graal, marques de structuration discursive et transitions narratives* », *L'Information grammaticale* 103, 29-36.

(2007) avec Heiden Serge et Lavrentiev Alexei. « Typologie des textes et des phénomènes linguistiques pour l'analyse du changement linguistique avec la Base de Français Médiéval », *Linx*, numéro spécial, 125 - 139.

(2007) avec Lavrentiev Alexei et Marchello-Nizia Christiane. « Les corpus de français médiéval : état des lieux et perspectives », *Revue française de linguistique appliquée* 121, 125-128.

(2008) avec Heiden Serge, Lavrentiev A. et Marchello-Nizia Christiane. « Constitution et exploitation des corpus d'ancien français et de moyen français », *Corpus* 7, 5-23.

(2012). « Le pronom anaphorique *cil* de l'ancien français : continuité ou discontinuité topicale ? », In : Denizot, C. & Dupraz, E. (éd.) *Anaphore et anaphoriques : variété des langues, variétés des emplois*, Mont Saint-Aignan : Publications des universités de Rouen et du Havre (Cahiers de l'ERAC), 97-115.

HEIDEN Serge

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: HEIDEN

Vorname: Serge

Alter : 45

Gegenwärtige Stellung: Ingénieur de Recherche en Informatique Linguistique à l'ENS de Lyon

Akademische Ausbildung und beruflicher Werdegang:

Docteur en informatique, Université Paris 6

- i. 1993-1994 : Attaché d'Enseignement et de Recherche en Linguistique Informatique à l'ENS de Fontenay/Saint-Cloud, Fontenay-aux-roses
- ii. 1994-2005 : Ingénieur d'Études en Informatique Linguistique à l'ENS de Fontenay/Saint-Cloud, Saint-Cloud ; puis ENS-LSH, Lyon
- iii. 2005- : Ingénieur de Recherche en Informatique Linguistique à l'ENS de Lyon

B/ Weitere berufliche Erfahrungen

I encadrement du projet ANR 2007-2010 Textométrie : Fédération des recherches et développements en textométrie autour de la création d'une plateforme logicielle ouverte

II depuis janvier 2007 : co-responsable de l'équipe S3CoDi, une des 3 équipes de l'UMR5191 ICAR - <http://icar.univ-lyon2.fr/Equipe3/index.htm>

III depuis mai 2010 : membre élu du Comité national du CNRS section 34

IV depuis mars 2010 : membre élu du conseil scientifique de l'ENS de Lyon

V organisation des JADT 2008 (Journées internationales d'Analyse statistique des Données Textuelles) - <http://jadt2008.ens-lyon.fr>

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

- **Heiden, S.** (2010). The TXM Platform : Building Open-Source Textual Analysis Software Compatible with the TEI Encoding Scheme. In Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation (PACLIC 24). 4-7 November 2010, Sendai.
- Guillot, C., **Heiden, S.**, Lavrentiev, A. (2005). Typologie des textes et des phénomènes linguistiques pour l'analyse du changement linguistique avec la Base de Français Médiéval. LINX, vol.(spécial – 2ième semestre) : 125-139.
- **Heiden, S.** (2004). Interface hypertextuelle à un espace de cooccurrences : implémentation dans Weblex. In « Le poids des mots », Purnelle, G., Fairon, C. and Dister, A. editors, Proc. of JADT'04 (7th Journées internationales d'Analyse Statistique des données Textuelles), pp. 577-588.
- **Heiden, S.**, Lavrentiev, A. (2004). Ressources électroniques pour l'étude des textes médiévaux : approches et outils. *Revue Française de Linguistique Appliquée*, vol.(9) : 91- 102.
- Prévost S., **Heiden S.** (2002). Etiquetage d'un corpus hétérogène de français médiéval : enjeux et modalités. In Pusch C. D. and Raible W., editors, *Romanistische Korpuslinguistik : Korpora und gesprochene Sprache, Romance Corpus Linguistics : Corpora and Spoken Language*, pp. 127-136.

LAVRENTIEV Alexei

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: Lavrentiev

Vorname: Alexei

Alter: 38

Gegenwärtige Stellung : Ingénieur de recherche CNRS, UMR ICAR

Akademische Ausbildung und beruflicher Werdegang:

- 2009 : Doctorat en sciences du langage, Thèse *Tendances de la ponctuation dans les manuscrits et incunables français en prose, du XIIIe au XVe siècle*, ENS LSH, Lyon ;

- 1999 : Candidat ès sciences philologiques (équivalent du doctorat du 3e cycle), thèse *Analyse typologique de la catégorie des cas en russe*, Université d'État de Tomsk (Russie) ;

- 1994 : D.E.A. de linguistique française, Université Paris 3 ;

- 1993 : Diplôme d'études supérieures (équivalent maîtrise) en langue et littérature russe, Université d'état de Novossibirsk.

B/ Weitere berufliche Erfahrungen

1998-2004 : Chercheur à l'Institut de philologie de la branche sibérienne de l'Académie des sciences de Russie ;

2001-2002 : Fulbright visiting fellow, Princeton University, participation au projet *Charrette* ;

2004-2008: Attaché temporaire d'enseignement de recherché, ENS LSH, Lyon, gestionnaire de la Base de Français Médiéval;

2008-2009 : Ingénieur de recherche contractuel, projets ANR CoRPTeF et ANR-DFG SRCMF.

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

- Lavrentiev, A. (2011) « Syntactic reference corpus of medieval French », Zakharov, V. (éd.), Proceedings of the International Conference “Corpus Linguistics – 2011”, St. Petersburg, 27-29 juin, St. Peterwbourg, SPBU, p. 41-45.
- Lavrentiev, A. (2010) « La 'phrase' en français médiéval : une réalité ou une reconstruction artificielle ? », Neveu, F., Muni Toke, V., Durand, J., Klinger, T., Mondada, L., Prévost, S. (éds), *Actes du 2e Congrès Mondial de Linguistique Française*, La Nouvelle Orléans, 12-15 juillet 2010, p. 277-289.
- Guillot, C., S. Heiden, A. Lavrentiev & Ch. Marchello-Nizia (2008) « Constitution et exploitation des corpus d'ancien et de moyen français », *Corpus*, n° 7, p. 5 – 23.
- Pincemin, B., C. Guillot, S. Heiden, A. Lavrentiev & Ch. Marchello-Nizia (2008) « Usages linguistiques de la textométrie. Analyse qualitative de la consultation de la Base de Français Médiéval via le logiciel Weblex », *Syntaxe & Sémantique*, n° 9, p. 87 – 110.
- Lavrentiev, A (2007) « Typologie textuelle pour l'étude linguistique de manuscrits français médiévaux », in Lavrentiev A. (dir.), *Systèmes graphiques de manuscrits médiévaux et incunables français : ponctuation, segmentation, graphies*. (Actes de la Journée d'étude de Lyon, ENS LSH, 6 juin 2005), Chambéry, Université de Savoie, p. 49 – 66.
-

LEEMAN Danielle

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: Leeman

Vorname: Danielle

Alter: 64

Gegenwärtige Stellung : Professeur des universités Paris-Nanterre

Akademische Ausbildung und beruflicher Werdegang:

Professeur des universités depuis 2001 à Paris-Nanterre. Enseignements de syntaxe, sémantique, linguistique générale aux divers niveaux de la Licence (Lettres classiques et modernes) et du Master professionnel de FLE.

Directeur de l'Ecole doctorale 139 *Connaissance, Langage, Modélisation* entre novembre 2003 et février 2009 (évaluation « A » en 2008).

Vice Présidente déléguée à l'International entre 2009 et 2011, et à nouveau à partir de mars 2012 au sein de l'Equipe nouvelle issue des élections présidentielles de l'université en 2012.

B/ Weitere berufliche Erfahrungen

Direction des revues internationales *Langue française* et *Langages* jusqu'en décembre 2009.

Deux mandats (4 ans) successifs, comme PR, au CNU (démission en décembre 2009 pour raisons personnelles).

Expertises pour le FNRS belge et pour le FNRS suisse (en moyenne deux par an depuis une dizaine d'années).

Expertises pour l'ANR (une en 2009 et une en 2010).
Membre du Comité d'évaluation AERES pour une UMR (2010) et trois EA (2011).
Présidente du Comité d'évaluation AERES pour trois ED du PRES de Lyon (2010).
Nommée par le Ministère de l'Education nationale comme membre de la Commission de suivi des enseignements de grammaire pour l'Ecole élémentaire et le collège (novembre 2011-...).

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

1. Leeman, D. dir. (2007) *Modèles linguistiques* 55 : *De la préposition à la locution prépositive*.
2. Leeman, D. dir. (2008) *L'Information grammaticale* 117 : *Locutions du français, traduction et traduction automatique*.
3. Leeman, D. dir. (2008) *Langue française* 157 : *Enigmatiques prépositions*.
4. Leeman, D. dir. (2010) *Le Français moderne* LXXVIII-1: *Sur sous et sur et leurs locutions*.
5. Leeman, D. (2011) « Complémentation verbale et préposition », in D. Van Raemdonck (dir.) *Modèles syntaxiques*, Berne, Peter Lang, coll. « Gramm'r » : 133-148.

PINCEMIN Bénédicte

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name : Pincemin

Vorname : Bénédicte

Alter : 41 ans

Situation professionnelle actuelle : Chargée de recherche CNRS

Akademische Ausbildung und beruflicher Werdegang :

1993 : Ingénieur de l'Ecole Centrale Paris.

1999 : Thèse de Doctorat en Linguistique, Paris IV Sorbonne, dir. François Rastier, dans le cadre d'un contrat avec EDF. Soutenue le 6 avril 1999, mention très honorable et félicitations. Intitulé : Diffusion ciblée automatique d'informations : conception et mise en oeuvre d'une linguistique textuelle pour la caractérisation des destinataires et des documents.

2000 : Admission au concours de recrutement de chargés de recherche 2e classe CNRS, section 34 (Représentations – langages – communications). 2004 : promotion au grade de chargée de recherche 1ère classe. 2000-2005 : affectation au LLI (Paris 13 Villetaneuse) ; depuis 2006 : affectation au laboratoire ICAR (ENS de Lyon).

B/ Weitere berufliche Erfahrungen

Gestionnaire de la liste électronique Sémantique des Textes (depuis 1995) (plus de 350 abonnés dans une vingtaine de pays) : <https://listes.cru.fr/sympa/info/sdt>

Co-éditeur du corpus XML-TEI Droits de l'Homme (2005, 2006) <http://www-ldi.univparis13.fr/> Projet ANR Fédération des recherches et développements en textométrie autour de la création d'une plateforme logicielle ouverte (ANR-06-CORP-029) : <http://textometrie.ens-lyon.fr/>: Coordination de l'équipe Conception et Application.

Membre du comité scientifique + formateur à l'Ecole thématique CNRS MISAT (Méthodes

Méthodes Informatiques et Statistiques en Analyse de Textes) (annuelle depuis 2009).

Membre du comité scientifique des Journées internationales d'Analyse statistique des Données Textuelles, co-organisatrice avec Serge Heiden de l'édition 2008 de la conférence à Lyon.

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

Heiden Serge, Magué Jean-Philippe, Pincemin Bénédicte (2010) - « TXM : Une plateforme logicielle open-source pour la textométrie -conception et développement », in Sergio Bolasco, Isabella Chiari, Luca Giuliano (eds), Statistical Analysis of Textual Data -Proceedings of 10th International Conference JADT 2010, Edizioni Universitarie di Lettere Economia Diritto, Rome, 9-11 juin 2010. [En ligne sur le site LED electronic archive : <http://www.ledonline.it>]

Pincemin Bénédicte, Heiden Serge, Lay Marie-Hélène, Leblanc Jean-Marc, Viprey Jean-Marie (2010) - « Fonctionnalités textométriques : Proposition de typologie selon un point de vue utilisateur », in Sergio Bolasco, Isabella Chiari, Luca Giuliano (eds), Statistical Analysis of Textual Data -Proceedings of 10th International Conference JADT 2010, Edizioni Universitarie di Lettere Economia Diritto, Rome, 9-11 juin 2010. [En ligne sur le site LED electronic archive : <http://www.ledonline.it>]

Pincemin Bénédicte, Guillot Céline, Heiden Serge, Lavrentiev Alexei, Marchello-Nizia Christiane (2008) - « Usages linguistiques de la textométrie. Analyse qualitative de la consultation de la Base de Français Médiéval via le logiciel Weblex. », Mathieu Valette (éd.), Textes, documents numériques, corpus. Pour une science des textes instrumentée, Syntaxe & Sémantique, 9, pp. 87-110.

Pincemin Bénédicte, Issac Fabrice, Chanove Marc, Mathieu-Colas Michel (2006) – « Concordanciers : thème et variations », Actes des 8es Journées internationales d'Analyse statistique des Données Textuelles (JADT 2006), Jean-Marie Viprey et al. (éds), Besançon : Presses Universitaires de Franche-Comté, ISBN 2.84867130.0, vol. II, pp. 773-784. [En ligne sur le site Lexicometrica : <http://www.cavi.univparis3.fr/lexicometrica/jadt/jadt2006/PDF/II-069.pdf>]

Pincemin Bénédicte (2004) - « Lexicométrie sur corpus étiquetés », Actes des 7es Journées internationales d'analyse statistique des données textuelles (JADT 2004), Gérald Purnelle & al. (éds), Presse universitaires de Louvain, Louvain-la-Neuve (Belgium), 10-12 mars 2004, vol. II, pp. 865-873. [En ligne sur le site Lexicometrica : http://www.cavi.univ-paris3.fr/lexicometrica/jadt/jadt2004/pdf/JADT_084.pdf]

SARDA Laure

A/ Name, Vorname, Alter, gegenwärtige Stellung, akademische Ausbildung und beruflicher Werdegang

Name: Sarda

Vorname: Laure

Alter: 39

Gegenwärtige Stellung: Chargée de recherche au CNRS (1er classe)

Akademische Ausbildung und beruflicher Werdegang:

2007-2010 Co-rédactrice en chef de la revue discours (<http://discours.revues.org/>)

2003 Mutation dans le laboratoire Lattice (CNRS/ENS, Paris)

- 2000 Séjour de recherche à l'université d'Austin, USA. (Collaboration avec Knud Lambrecht).
- 1999- Chargée de recherche au CNRS (à l'ERSS, UMR 5610, Toulouse).
- 1999 Thèse de doctorat en linguistique générale et théorique (avec les félicitations)
Université Toulouse le Mirail
- 1994 DEA en linguistique générale et théorique (avec les félicitations)
Université Toulouse le Mirail
- 1993 Maîtrise en linguistique générale et théorique (avec les félicitations)
Université Toulouse le Mirail
- 1992 Licence en linguistique générale et théorique
Université Toulouse le Mirail
- 1991 DEUG en Lettres Modernes
Université de Perpignan

B/ Weitere berufliche Erfahrungen

Enseignement de syntaxe et sémantique en deug et licence de Sciences du Langage à l'université de Toulouse-le-Mirail, entre 1994 et 2002.

C/ Fünf Veröffentlichungen in Zusammenhang mit dem Thema des Forschungsprojektes

- Sarda L. & Charolles M. (eds) (soumis) Les adverbiaux prépositionnels : position, fonction et portée. Numéro thématique, *Travaux de Linguistique*.
- Carlier A. & Sarda L. (2010). Le complément de la localisation spatiale : entre argument et adjectif, In Franck Neveu, Valelia Muni Toke, Thomas Klingler, Jacques Durand, Lorenz Mondada et Sophie Prévost (Eds.) *CMLF 2010 - 2ème Congrès Mondial de Linguistique Française*, Publié en ligne : 12 Juillet 2010, CMLF 2010, 138.
- Carter-Thomas, S. Sarda, L. (2011). Identifying the semantic roles of postverbal prepositional phrases: the participant/circumstance interface, *La linguistique systémique fonctionnelle et la langue française* (Ed D. Banks).
- Sarda Laure (2010). "Les adverbiaux prépositionnels en *dans* : exploration en corpus de la notion de contenance ". *Corela*, Numéros spéciaux, Espace, Préposition, Cognition-Hommage à Claude Vandeloise. <http://edel.univ-poitiers.fr/corela/document.php?id=2645>
- Sarda, L., Carter-Thomas, S. (2010). Rôle informationnel et textuel des adverbiaux détachés : exemples anglais et français en *sur* et *on*. In *Directions actuelles en linguistique du texte*. L. Florea, C. Papahagi, L. Pop, A. Curea (eds) Actes du colloque international *Le texte : modèles, méthodes, perspectives* Casa Cărții de Știință, Cluj-Napoca : 255-268.
- Fagard B. & Sarda L. 2009. « Etude diachronique de la préposition dans ». *Autour de la préposition*, J. François, E. Gilbert, C. Guimier & M. Krause (dir.), Bibliothèque de Syntaxe et sémantique, PUC, Caen : 225-237.

Fiche de poste Projet ANR-DFG PRESTO : profil ingénieur informatique partenaire français

Intitulé du poste : Concepteur-développeur informatique open-source en traitement linguistique automatique (TAL)

Service de rattachement : Laboratoire ICAR UMR 5191 au sein de l'équipe ICAR3, ENS de Lyon

Place du poste dans l'organigramme du service : Sous la responsabilité directe du porteur français du projet

Missions principales

- ▶ Contribuer à la formalisation d'un ensemble de besoins en TAL à partir de rencontres avec les chercheurs.
- ▶ Réaliser un travail de veille sur les technologies logicielles recrutables (en tenant compte de leur licence de diffusion) pour le développement informatique du projet
- ▶ Ecrire le programme source correspondant au modèle de conception selon les règles de l'art du développement open-source en tenant compte des technologies open-source recrutables appropriées
- ▶ Vérifier la conformité du logiciel développé au modèle et aux besoins exprimés (recette)
- ▶ Participer à la diffusion du logiciel (mise en ligne du source et du binaire, suivi de bugs, formations, etc.) et à son évolution (suivi open-source)
- ▶ Participer à l'utilisation du logiciel pour la préparation du corpus

Spécificités liées au poste : autonomie, obligations, responsabilités, contexte, particularités

- ▶ L'ingénieur établit son plan de travail en concertation avec le porteur de projet et l'exécute de façon autonome
- ▶ Le cas échéant, il travaille en coordination avec d'autres compétences participant au projet (concepteurs, développeurs, etc.)
- ▶ En cas de difficulté à atteindre ses objectifs ou à respecter les délais prévus, il doit en informer le porteur du projet

Champ des relations internes et externes à l'établissement

- ▶ L'ingénieur est en contact permanent avec les partenaires du projet en fonction de leurs responsabilités respectives au sein du projet.
- ▶ Ses horaires de travail et ses congés correspondent à ceux de la catégorie correspondante au sein de l'ENS de Lyon

Connaissances et compétences

Connaissances : expérience en conception et développement Java open-source pour le TAL.

Compétences :

Requises

- ▶ Maîtriser le langage Java et une méthode de conception associée
- ▶ Connaissance des langages de script Groovy et R
- ▶ Environnement de développement open-source : IDE Eclipse, SVN, Javadoc, etc.

▶ Système d'exploitation : Linux

▶ Anglais courant

Désirées

▶ Maîtrise du traitement des formats XML (TEI P5)

- ▶ Connaissance d'outils de tokenization de texte brut en français
 - ▶ Connaissance d'outils d'étiquetage morphosyntaxique et de lemmatisation du français
 - ▶ Connaissance d'outils de chunking ou d'analyse syntaxique du français
 - ▶ Connaissance de modèles statistiques appliqués aux données linguistiques
-

Fiche de poste Projet ANR-DFG PRESTO : profil post-doctorant-e- partenaire français

Localisation : ENS Lyon, Laboratoire ICAR (Equipe ICAR 3)

Durée du contrat : CDD de 18 mois

Contexte du projet

Etude de l'évolution, sur corpus, en diachronie longue (IX^e-XX^e s) du profil combinatoire de quelques prépositions françaises

Le projet PRESTO a pour but l'étude diachronique de la combinatoire des prépositions françaises à, en, par, contre, dès, devant, entre, pour, sans, sur, sous, vers, dans, de l'ancienne langue (IX^e siècle) jusqu'au français contemporain. Adossée à une analyse distributionnelle à l'interface entre (morpho-)syntaxe et sémantique, elle sera menée sur la base d'études statistiques du comportement des prépositions dans différents genres au sein d'un corpus construit pour les besoins du projet.

Tâches

Il s'agira pour le (/ la) chercheur (-se) :

- ▶ de participer étroitement à l'étude pilote du projet, consistant à examiner l'évolution du profil combinatoire (valence syntaxique, identification et hiérarchisation des accompagnateurs spécifiques, calcul d'indices de similarité, échelonnement multidimensionnel, ...) des prépositions « en » et « dans » au cours de la période XVII^e-XX^e s, en recourant aux outils statistiques et textométriques (plateforme TXM) mis en œuvre dans le projet.
- ▶ De participer à l'exploitation et à l'interprétation des résultats quantitatifs obtenus en vue d'étudier, dans le corpus XVII^e-XX^e i) l'émergence, la réorganisation et le déclin de structures prépositionnelles figées ii) d'identifier des schémas co-occurentiels se manifestant dans des unités textuelles plus vastes (cadres et séquences de cadres de discours) et d'étudier leur évolution.
- ▶ De participer à l'enrichissement du corpus pour la tranche temporelle XVII^e- XX^e s en vue d'équilibrer l'échantillonnage (par tranches de siècle et par genres).
De participer à l'annotation du corpus avec les différentes informations linguistiques nécessaires au projet (segmentation lexicale, morphosyntaxe, lemme, structures syntaxiques, groupes prépositionnels et leurs fonctions syntaxiques)
- ▶ De participer activement aux discussions préalables aux différentes phases d'annotation du corpus (annotations morphosyntaxiques et syntaxiques, codage des positions des GP dans différentes unités compositionnelles (phrase, paragraphe, ...), typologie des genres discursifs et des types de texte adoptée sur la base des descripteurs textuels utilisés dans la BFM (domaine, genre, forme, etc.), ...)

Profil du / de la candidat-e

Le/la candidat-e est titulaire d'une thèse de Sciences du Langage. Il est souhaitable qu'il /elle ait consacré un volet de sa recherche doctorale ou quelques-uns de ses travaux à l'étude de certaines prépositions et/ou groupes prépositionnels. Il/elle possède des connaissances solides dans le domaine de la diachronie et a mené des travaux (éventuellement publiés) s'appuyant sur la théorie de la grammaticalisation. Il / elle doit par ailleurs pouvoir témoigner d'une pratique de recherche adossée aux corpus (thèse, publications, communications) et se montrer intéressé(e) par l'acquisition des méthodes informatiques (préparation de corpus) et statistiques utilisées en analyse de textes (outils développés à Cologne – partenaire allemand - et outils textométriques développés à Lyon). La maîtrise préalable de tels outils serait un plus.

II. 2. Allgemeine Bibliographie

- Amiot, D. & W. De Mulder, (2002), De l'adverbe au préfixe en passant par la préposition : un phénomène de grammaticalisation ?, *Linguisticae Investigationes* XXV/2, 247-273.
- Bekkers, F. (1931), *Etude sur l'emploi que Froissart fait de la préposition*. Dissertation, Groningen.
- Blumenthal, P. (2002), Profil combinatoire des noms : synonymie distinctive et analyse contrastive, *Zeitschrift für französische Sprache und Literatur* 112 : 115-138.
- Blumenthal, P. (2003), Der Begriff der externen und internen Sprachgeschichte in der Romanistik, in *Romanische Sprachgeschichte* 1. Berlin :Teilband, 38-45.
- Blumenthal, P. (2006), *Wortprofil im Französischen* (Beihefte zur ZrP 332), Tübingen.
- Blumenthal, P. (2007), Histoire de la montée en puissance de *début*, in: L. Begioni/C. Muller (s.l.d.): *Problèmes de sémantique et de syntaxe - Hommage à André Rousseau*, Lille : Editions du Conseil Scientifique de l'Université Charles-de-Gaulle - Lille 3, (Travaux et recherches), 445-471.
- Blumenthal, P. (2008), Combinatoire des prépositions : approche quantitative, *Langue Française*, 157, 37-51.
- Blumenthal, P. (2009). "Les noms d'émotion : trois systèmes d'ordre", in: I. Novakova & A. Tutin (eds), *Le Lexique des émotions*, Grenoble : ELLUG 2009 (Langues, Gestes, Paroles), 41-64.
- Blumenthal, P. (2011), Le figement : du XVII^e siècle au français contemporain, in: J.-C. Anscombre/S. Mejri (éds.): *Le figement linguistique: la parole entravée*, Paris: Éditions Champion 2011 (Lexica 18), 283-302.
- Blumenthal, P. (2012, à paraître) : Évolutions de la combinatoire prépositionnelle : le cas de *en*, *Mélanges Olivier Soutet*.
- Blumenthal P., S. Diwersy & J. Mielebacher (2005), Kombinatorische Profile und Profilkontraste: Berechnungsverfahren und Anwendungen, *Zeitschrift für Romanische Philologie* n°121, 1: 49-83.
- Blumenthal, P. & M.-J. Hausmann (2006), *Collocations, corpus, dictionnaires*, *Langue Française*, 150.
- Blumenthal, P. & S. Mejri (2008) (eds), *Les séquences figées : entre langue et discours*, *Zeitschrift für Französische Sprache und Literatur*, Beihefte 36, Franz Steiner Verlag : Stuttgart.
- Bonami, O. (1999), *Les constructions du verbe : le cas des groupes prépositionnels argumentaux*, Thèse de doctorat, Université Paris 7.
- Bolly, C. (2008), *Les unités phraséologiques : un phénomène linguistique complexe ?* Thèse de doctorat, Université Louvain-la-Neuve.
- Buridant, C. (2000), *Grammaire Nouvelle De L'ancien Français*. Paris : Sedes.
- Butler, C. (2003), *Structure and Function: A Guide to Three Major Structural-Functional Theorie*, Amsterdam, Benjamins.

- Cadiot, P. (1997a), Les paramètres de la notion de préposition incolore, *Faits de langues*, 9, 127-134
- Cadiot, P. (1997b), *Les prépositions abstraites en français*, Paris : Armand Colin.
- Carlier A. & Sarda L. (2010). Le complément de la localisation spatiale : entre argument et adjectif, In F. Neveu, V. Muni Toke, T. Klingler, J. Durand, L. Mondada & S. Prévost (eds.) *CMLF 2010 - 2ème Congrès Mondial de Linguistique Française*.
- Cervoni, J. (1991), *La préposition*, Paris-Louvain-la-Neuve, Duculot.
- Charolles, M. (1997), *L'encadrement du discours : univers, champs, domaines et espaces*, Cahier de Recherche Linguistique, LANDISCO, URA-CNRS 1035 Université Nancy 2, n° 6, 1-73.
- Charolles, M. & S. Prévost (2003)(eds.), *Adverbiaux et topiques. Travaux de linguistique* 47.
- Charolles, M. & M.-P. Péry-Woodley (2005)(eds.), *Les adverbiaux cadratifs. Langue française* 148.
- Charolles, M., Vigier, D. (2005), Les adverbiaux en position préverbale : portée cadrative et organisation des discours. *Langue Française*, 148, 9-30.
- Coseriu, E. (1974), *Synchronie, Diachronie und Geschichte. Das Problem des Sprachwandels*, München.
- Croft, W. & A. Cruse (2004), *Cognitive Linguistics*, Cambridge, CUP.
- Dannell, K. J., Nothing but phrases. About the distribution of idioms and stock phrases, in Edlund, L.E. & Persson, G. (éds). *Language: the time machine*, Umeå : Umeå University
- Dendale, P. & W. De Mulder, (1998), *Contre et sur : du spatial au métaphorique ou inversement ?*, *Verbum* XX/4, 405-434.
- De Mulder, W. & D. Stosic (2009) *Présentation*, in De Mulder, W. & Stosic, D. (2009) (eds) *Approches récentes de la préposition, Langages*, 173, 3-13.
- De Mulder, W. & A. Vanderheyden (2002), The evolution of French 'sur' : steps toward a diachronic prototype semantics of spatial prepositions, In Hubert Cuyckens (ed.), *Perspectives on prepositions*. Tübingen : Niemeyer, 177-193.
- Denis, P., Sagot B. (2009), « Coupling an annotated corpus and a morphosyntactic lexicon for state-of-the-art POS tagging with less human effort », in *Proceedings of The Pacific Asia Conference on Language, Information and Computation (PACLIC 23)*, Hong Kong, China.
- Dufresne, M., F. Dupuis & M. Tremblay (2003), Les prépositions dans l'histoire du français : transitivité, grammaticalisation et lexicalisation, *Verbum* XXV/4, 551-564.
- Everaert, M. et al.(eds). 1995. *Idioms*. Hillsdale, N.J.: Erlbaum.
- Fagard, B. (2004), *Après : de l'espace au temps, la sémantique en diachronie*. In A. Soares da Silva, A. Torres & M. Gonçalves (eds.), *Linguagem, cultura e cognição*, Braga : Almedina, 231-246.
- Fagard, B. (2006), La grammaticalisation en question – du latin aux langues romanes modernes, *Modèles linguistiques* XXVII-1/53, 91-110.
- Fagard, B. (2007), *De puis à depuis : préfixation et évolution sémantique*. In B. Combettes & C. Marchello-Nizia (eds.), *Etudes sur le changement linguistique du français*, 129-143.
- Fagard, B. (2008), 'Côté' dégrammaticalisation – le cas des prépositions. In B. Fagard, S. Prévost, B. Combettes & O. Bertrand (eds.), *Evolutions en français*, Berne : Peter Lang.
- Fagard, B. 2010. *Espace et grammaticalisation - L'évolution sémantique des prépositions dans les langues romanes*. Sarrebruck : Editions Universitaires Européennes.
- Fagard, B. & W. De Mulder (2007), La formation des prépositions complexes : grammaticalisation ou lexicalisation ?, *Langue française* 156, 9-29.
- Fagard, B. & L. Sarda (2009), Etude diachronique de la préposition *dans*, in J. François, É. Gilbert, C. Guimier & M. Krause (eds.), *Autour de la préposition*, Caen : Presses Universitaires de Caen, 225-236.

- Fournier, N. (1998), *Grammaire du français classique*, Paris : Belin
- Franckel J.J. & Paillard D. (2006), De la couleur des prépositions dans leurs emplois fonctionnels, *Modèles Linguistiques*, XXVII, 2, 54, 51-56.
- Franckel J.J. & Paillard D. (2008), *Grammaire des prépositions*, Paris : Ophrys.
- Gougenheim, G. (1974), *Grammaire de la langue française du 16e siècle*. Paris : Picard.
- Gross, G. (1996), *Les expressions figées en français. Noms composés et autres locutions*, Paris : Ophrys.
- Grossmann, F. & A.Tutin (2003) (eds), *Les collocations. Analyse et traitement*, Amsterdam : De Werelt.
- Grunig, B.-N. (1997), La locution comme défi aux théories linguistiques : une solution d'ordre mémoriel?, in Martins-Baltar, M. (éd.), *La locution entre langue et usages*, Paris Fontenay-St Cloud : ENS, 225-240.
- Heine B., U. Claudi & F. Hünemeyer (1991), *Grammaticalization : A Conceptual Framework*. Chicago : University of Chicago Press.
- Heine, B. & T. Kuteva, (2002), *World Lexicon of Grammaticalization*. Cambridge, Cambridge University Press.
- Hopper P. & E. Traugott (éds) (1993), *Grammaticalization*. Cambridge University Press.
- Hoey, M. (2005), *Lexical Priming*, London, Routledge
- Homles, M. (2008) « Domestic strife in early modern Europe: images and texts in a virtual anthology ». (Conference presentation) with Claire Carlin; Digital Humanities 2008.
- Homles, M. (2010) « Using the Universal Similarity Metric to Map Correspondences between Witnesses ». (Conference presentation) Digital Humanities 2010. Conference, Kings College London..
- Hudson, J. (1998), *Perspectives on fixedness: applied and theoretical*, Lund Studies in English, 94. Lund: Lund University Press.
- Hunston, S. & G. Francis, (2000), *Pattern Grammar*, Amsterdam: Benjamins.
- Kilroe, P. (1989), *The grammaticalization of French à*. Doctoral dissertation, University of Texas : Austin.
- Kilroe, P. (1994), The grammaticalization of French à. In William Pagliuca (ed.), *Perspectives on Grammaticalization*, Amsterdam / Philadelphia : J. Benjamins, 49-61.
- Lamiroy, B., (2010), *Les expressions verbales figées de la francophonie. Les variétés de Belgique, de France, du Québec et de Suisse*. Paris :Ophrys.
- Lamiroy, B. & J. Klein (2010), Le problème central du figement est le semi-figement, in *Linx*, 53 (2005), 135-154.
- Lang, J. (1991), *Die französischen Präpositionen. Funktion und Bedeutung*, Heidelberg.
- Lavieu, B. (2006), Distinction entre groupes prépositionnels régis et non régis. *Modèles linguistiques* 53, tome XXVII-1 : 131-144.
- Lebel, M-É. (2003), *Le morphème par en ancien français*. Thèse en linguistique, Université du Québec à Montréal
- Lecolle, M. & S. Leroy (2006) (eds), *Changements linguistiques : figement, lexicalisation, grammaticalisation*, Cahiers de praxématique, 46 .
- Le Draoulec, A. & Vigier, D. (2009). Dans la nuit je rentre chez moi. Interprétations temporelle et spatiale de quelques constructions dans + SNtps. Actes des cinquièmes Rencontres de Sémantique et Pragmatique - « Espace, Temps dans les langues », *Revue de sémantique et de pragmatique*.
- Leeman, D. (1998), *Les circonstants en question(s)*, Paris, Kimé.
- Leeman, D. (1999), L'unité lexicale dans la perspective harrissienne, *LINX* 40 : 117-136.
- Leeman, D. (2006), la préposition française : caractérisation syntaxique de la catégorie, in Leeman, D. (2006) (éd), *la préposition en français, I*, *Modèles linguistiques* 54, tome XXVII, 7-17.
- Leeman, D. (2008), Prépositions du français, état des lieux, in Leeman, D. (2008) (éd), *Enigmatiques prépositions*, *Langue Française*, 157, 5-19.
- Le Pesant, D. (2006), Classification à partir des propriétés syntaxiques, *Modèles linguistiques* 53, tome XXVII-1 : 51-74.

- Marchello-Nizia, C. (1997 [1979]), *Histoire de la langue française aux 14^{ème} et 15^{ème} siècles*. Paris : Bordas.
- Marchello-Nizia, C. (1999), *Le Français en diachronie*, Paris : Ophrys.
- Marchello-Nizia, C. (2002), Prépositions françaises en diachronie : une catégorie en question, *Linguisticae Investigationes* XXV/2, 205-221.
- Martineau, F. (2000), Les prépositions à et de de l'ancien français au français classique. In Claude Buridant (dir.), *Actes du IX^e colloque international sur le moyen français*. Strasbourg : Presses Universitaires de Strasbourg, 165-183.
- Marchello-Nizia, C. (2006), *Grammaticalisation et changement linguistique*. Bruxelles : Duculot / De Boeck
- Mejri, S. (2000), Figement lexical et renouvellement du lexique : quand le processus détermine la dynamique du système, *Le français moderne* LXVIII (1), 39-62.
- Melis, L. (2003), *La préposition en français*, Paris : Ophrys.
- Mel'čuk, I. (2003), Collocations : définition, rôle et utilité, In F.Grossman et A.Tutin (éds), 23-31.
- Michael, F. (1904), *Über den Gebrauch der Präpositionen mit Ausschluss von a, de, en, im altfranzösischen Rolandsliede*. Dissertation, Greifswald.
- Moon, R. (1998), *Fixed Expressions and Idioms in English, a Corpus-Based Approach*, Oxford : Clarendon
- Noailly, M. (2010), L'adjectif, quand il est invariable, sort-il pour autant de sa catégorie ?, in J. Goes & E. Moliné (eds), *L'adjectif hors de sa catégorie*, Actes du sixième colloque franco-roumain, Université d'Artois, 23-25 mai 2007, Arras.
- Raithel, G. (1875), *Über den Gebrauch und die begriffliche Entwicklung der altfranzösischen Präpositionen od, par, en*. Dissertation, Göttingen.
- Sagot B. (2010), « The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French », in *Proceedings of the 7th Language Resources and Evaluation Conference (LREC'10)*, Vallette, Malta.
- Sarda L. & Charolles M. (eds) (soumis) Les adverbiaux prépositionnels : position, fonction et portée. Numéro thématique, *Travaux de Linguistique*.
- Senellart, J. (1998), Reconnaissance automatique des entrées du lexique-grammaire des phrases figées, in Lamiroy, B. (éd), *Le lexique-grammaire. Travaux de Linguistique*, 37, numéro spécial. 109-127.
- Shears, F. (1922), *Recherches sur les prépositions dans la prose du moyen français (XIV^e et XV^e siècles)*. Paris : Champion.
- Shyldkrot, H. B.-Z. & S. Kemmer (1995), La grammaticalisation des prépositions : concurrence et substitutions, *Revue Romane* 30, 205-226.
- Stein, A. (2003), « Étiquetage morphologique et Lemmatisation de textes d'Ancien Français. » in: *Ancien et Moyen Français sur le Web: enjeux méthodologiques et Analyse du Discours*, P. Kunstmann & al. (eds), Ottawa: David Éditions, 273-284.
- Stein, A. and al. (eds) (2006), *Nouveau Corpus d'Amsterdam. Corpus informatique de textes littéraires d'Ancien Français (1150-1350)*, texte établi par A. Dees (Amsterdam 1987), remanié par A. Stein, P. Kunstmann & M.-D. Glessgen. Stuttgart: Institut für Linguistik/Romanistik.
- Svensson, M.H. (2004), Critères de figement. L'identification des expressions figées en français contemporain. Umeå : Umeå University.
- Thompson, S. (1985), Grammar and written discourse: Initial vs. final purpose clause in English, *Text* 5(1-2), 55-84.
- Traugott E.C.. & König E. (1991), The Semantics-Pragmatics of Grammaticalization Revisited, in E.C. Traugott & B. Heine (eds), *Approaches to Grammaticalization*, vol. 1, Amsterdam / Philadelphia, John Benjamins, 189- 218.
- Vaguer, C. (2006), Bibliographie générale des prépositions du français (I et II). *Modèles linguistiques* 54, tome XXVII-II : 171-203.
- Vaguer, C. (2007), Bibliographie générale des prépositions du français (III). *Modèles linguistiques* 55, tome XVIII-I.

- Vaguer, C. (2008), Classement syntaxique des prépositions simples du français, *Langue Française*, 157, 20-36.
- Vigier, D. (2003), Les syntagmes prépositionnels en « en N » détachés en tête de phrase référant à des domaines d'activité, *Linguisticae Investigationes*, 26 (1), 97 - 122, Amsterdam, John Benjamins B.V.
- Vigier, D. (2004), *Les groupes prépositionnels en « En N » : de la phrase au discours*, Thèse de 3^è cycle, Université Paris 3.
- Vigier, D. (2005). Les adverbiaux praxéologiques détachés en position initiale et leur portée. *Verbum*, XXVII, 2005, 3, 293-312.
- Vigier, D. (2008a). Les adverbiaux de localisation spéculaire : détachement et portée. In D. Apothéloz, B. Combettes, F. Neveu (eds) *Les linguistiques du détachement. Actes du colloque international de Nancy (7-9 juin 2006)*. Berne : P. Lang, 567-580.
- Vigier, D. (2008b), Contribution à une étude des constructions antéposées du type : *En homme intelligent et humain, il partagea tout de suite l'inquiétude de Marcel* (J. Verne), *Discours*, 2.
- Vigier D. (2008c). La gestion des cadres de discours dans une tâche rédactionnelle en FLE. In H. Hilton (ed). *Acquisition et didactique 1. Actes de l'atelier didactique, colloque international AFLS 2005*. Chambéry: Presses Universitaires de Chambéry, 113-129.
- Vigier, D. (à par., 2012a), « Comportements, déguisements, rôles, ... De quelques emplois intraprédicatifs de *en* », *Linguisticae Investigationes*, Amsterdam, Philadelphia : John Benjamins Publishing Co.
- Vigier, D. (à par., 2012b), « *En attendant* : un cas de pragmatization », in *Les adverbiaux prépositionnels : position, fonction et portée*, Charolles, M. & Sarda, L. (éds.), *Travaux de linguistique*.
- Vigier, D. (ed.)(en prép.), *La préposition « en »*, *Langue Française*, paris : Hachette
- Vigier, D., Terran E. (eds) (2005). *Les adverbiaux cadratifs et l'organisation des textes*. *Verbum*, XXVII, 3.
- Virtanen, T. (1992a.), *Discourse functions of adverbial placement in English*. Åbo: Åbo Akademi University Press.
- Virtanen, T. (1992b), Temporal Adverbials in Text Structuring: On Temporal Text Strategy. In A.-C. Lindberg, N. Enkvist & K. Wikberg (eds.), *Nordic Research on Text and Discourse*. Åbo: Åbo Academic Press, 85-197.
- Wray, A. (2002), *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.