



HAL
open science

Le CRISCO de Caen et l'investigation des interfaces linguistiques

Jacques François

► **To cite this version:**

Jacques François. Le CRISCO de Caen et l'investigation des interfaces linguistiques. *Zeitschrift für Französische Sprache und Literatur*, 2004, 114 (3), pp.267-281. <hal-00012468>

HAL Id: hal-00012468

<https://hal.science/hal-00012468v1>

Submitted on 24 Oct 2005

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

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



HAL Authorization

Le CRISCO de Caen ou les interfaces au premier plan

par
Jacques François
Directeur du CRISCO

Le *Centre de Recherches Interlangues sur la Signification en Contexte* est une équipe jeune, rattachée à l'*Institut de Linguistique Française*¹, qui a vu le jour en 2000 à l'Université de Caen sous l'impulsion de Jacques François — son actuel directeur assisté de Franck Neveu — de Claude Guimier, directeur de l'ELSAP, équipe du CNRS dont il est partiellement issu, et de Nicole Le Querler. L'intitulé de l'équipe pointe sur deux des interfaces qui focalisent l'intérêt de ses membres, celle entre langues de groupes divers (français moderne et ancien, anglais, allemand, italien, finnois, basque) et celle entre signification en langue et effets de sens en contexte. L'étude de cet espace de variation sémantique met en jeu un troisième type d'interface, celle entre syntaxe et sémantique, ce qui a donné le nom de la revue *Syntaxe et sémantique* éditée par l'équipe. Enfin une partie des travaux de linguistique fondamentale de l'équipe s'adossent sur deux outils de traitement automatique développés par l'équipe, le logiciel de synthèse vocale *KALI* pour le français et l'anglais et le *Dictionnaire Electronique des Synonymes* (abrégé *D.E.S.*).

1. Le CRISCO dans le paysage des sciences du langage en France

1.1. La place que s'est faite le CRISCO dans le paysage des sciences du langage en France tient d'une part à son activité de publication et de diffusion de travaux en rapport avec les interfaces évoquées plus haut et d'animation de la recherche linguistique dans l'espace européen et méditerranéen, mais surtout au développement et aux recherches associées au *D.E.S.* et au synthétiseur *KALI*.

La revue *Syntaxe et sémantique* éditée par J. François & N. Le Querler depuis 2000 accueille en priorité des numéros thématiques assez épais consacrés aux structures syntaxiques et sémantiques, spécialement dans leurs rapports réciproques. Après deux premiers numéros consacrés respectivement à la syntaxe des connecteurs dans différentes langues (cf. Guimier, coord. 2000) et à la sémantique du lexique verbal d'un point de vue linguistique, psychologique et en traitement automatique (cf. Cordier, Victorri & François 2001), la revue s'est tournée plus récemment vers les mots-outils du français, la valence verbale et adjectivale et l'articulation entre la polysémie lexicale et le statut de « polylexème ». A la revue s'adosse la *Bibliothèque de Syntaxe & Sémantique* dont le second volume est consacré aux Actes du colloque *L'adjectif en français et à travers les langues* organisé par le CRISCO en 2001 (cf. François, coord. sous presse).

¹ Fédération des équipes du CNRS orientées exclusivement ou majoritairement vers la linguistique du français.

La diffusion des rapports de recherche du laboratoire passe par les *Cahiers du CRISCO* dont les numéros récents sont téléchargeables sur le site du CRISCO. Plusieurs de ces cahiers donnent une image vivante de manifestations scientifiques du laboratoire².

1.2. Le *D.E.S.* possède l'originalité d'être en consultation gratuite sur Internet ; il possède 49000 entrées, reliées entre elles par 200000 relations synonymiques symétriques. Sur le site Web du CRISCO (<http://www.crisco.unicaen.fr>), il est possible d'obtenir les synonymes d'un mot demandé, et de naviguer ensuite dans le dictionnaire, puisque sur la page de réponse, chaque mot présent dans la liste des synonymes est lui-même un lien vers sa liste de synonymes. Cette hypertextualité explique que le *D.E.S.* connaisse un grand succès : en février 2004, il a reçu plus de 2 millions de requêtes. Son succès s'explique aussi par la convivialité de son interface : le programme de consultation corrige en effet toutes les fautes d'accents, redirige les formes fléchies vers la forme canonique des mots-vedettes, et tolère les fautes d'orthographe les plus courantes. De plus, il s'enrichit constamment de nouvelles entrées, choisies en analysant les requêtes restées sans réponse : les plus demandées seront insérées (dans la mesure où l'on peut leur donner des synonymes) dans le *D.E.S.*. En outre, le *D.E.S.* donne accès sur le Web aux représentations du sens qui ont été développées au laboratoire ; d'une part, ces représentations interactives donnent aux utilisateurs la possibilité d'explorer et de découvrir l'espace sémantique des mots, et d'autre part, une lettre d'information mensuelle gratuite diffuse aux abonnés non seulement les indications concernant l'emploi de ces représentations, mais aussi apporte les éléments nécessaires à leur compréhension en détaillant les principes du modèle sous-jacent. En outre, le *D.E.S.* et ses représentations ouvrent des perspectives pédagogiques et méthodologiques que nous parcourons dans divers projets qui concernent la dictionnaire, la lexicologie ou encore la psycholinguistique³.

1.3. La synthèse vocale est une technique qui permet de faire prononcer de façon intelligible et aussi naturelle que possible un texte écrit. Les synthétiseurs vocaux peuvent être utilisés dans l'industrie (pour "faire parler" des machines et automates), dans le secteur du multimédia, dans des serveurs vocaux consultables par téléphone, et bien sûr dans le domaine du handicap visuel ou vocal. Le synthétiseur vocal est un outil très important pour une personne aveugle : il permet à celle-ci l'accès aux données (lecture, bloc-notes, services de messagerie, utilisation de l'informatique). Le projet *KALI*, grâce à la collaboration d'utilisateurs déficients visuels, est conçu pour être parfaitement adapté au handicap visuel, mais la qualité obtenue nous permet de toucher les marchés grand public. Après trois ans d'études, le synthétiseur *KALI* a été commercialisé en 1999, avec deux voix masculines et une voix féminine en français. L'anglais britannique est en cours de développement. Une

² Cf. sur la contribution des unités lexicales et grammaticales à l'unité du discours le Cahier 12 (Legallois, coord. 2003), sur l'exploitation du *D.E.S.* combiné avec la base de données textuelles catégorisée FRANTEXT le Cahier 13 (François, Manguin & Victorri 2003) et sur la théorie « structurale-fonctionnelle » de la *Role and Reference Grammar* le Cahier 14 (François, coord. 2003).

³ Le *D.E.S.* est développé en collaboration étroite avec la société MEMODATA spécialisée en sémantique computationnelle. Dans la conception du *Dictionnaire Intégral* produit par MEMODATA, (i) le cœur du dictionnaire est une architecture qui doit permettre de décomposer les significations des mots en différentes parties, et en retour, à partir d'une décomposition, de rétablir une signification, il s'agit donc d'une sorte d'étude de la paraphrase appliquée à la définition elle-même ; (ii) MEMODATA est parti de divers courants de théories de la décomposition lexicale, pour l'essentiel la théorie sens-texte, la sémantique componentielle, le repérage notionnel, et cherche à les combiner de manière à atteindre le but recherché ; (iii) du fait qu'évidemment ce n'est pas tant le dictionnaire qu'il faut comprendre mais bien les occurrences réelles d'énoncés, MEMODATA effectue de nombreuses opérations appliquées au texte, en utilisant la ressource générale, à la fois pour valider cette ressource et pour offrir des outils d'ingénierie linguistique (cf. Dutoit, Nugues & de Torcy 2003 ; Dutoit & Poibeau, à paraître).

démonstration interactive est disponible sur le site www.crisco.unicaen.fr. L'architecture de KALI comporte six modules que traverse successivement le texte à prononcer (cf. Morel & Lacheret-Dujour 2001) :

► *Prétraitement* : Cette opération a pour but de faciliter la mise en oeuvre des différents traitements, notamment en développant les abréviations et sigles.

► *Analyse syntaxique et pragmatique* : L'analyse syntaxique fournit un découpage du texte en groupes intonosyntaxiques ainsi que leurs relations de dépendance. Cette opération est la base de la génération automatique de prosodie, même si elle doit être complétée par une analyse pragmatique du texte (focalisation, attitudes et plus généralement structure communicative)

► *Transcription graphème-phonème* : Pour prononcer correctement un texte, il faut connaître de nombreuses règles de prononciation. Le module de transcription est chargé de fournir un texte phonémique à partir du texte alphabétique. Le français comporte environ mille règles, l'anglais plusieurs milliers. Les mots d'emprunt et les noms propres nécessitent plusieurs milliers de règles supplémentaires (cf. Morel & Lacheret-Dujour 1998).

► *Base de diphones* : A partir de la voix d'un locuteur, on extrait une base de diphones, c'est-à-dire un peu plus de 1000 segments de signal qui seront ensuite concaténés pour former le signal acoustique de parole. Chaque segment va de la moitié d'un phonème à la moitié du suivant. L'interface entre les segments est le phonème, ce qui limite le nombre de segments nécessaires à la fabrication de n'importe quel énoncé. Les diphones étant saisis dans des contextes différents, il est nécessaire de les modifier pour que leur raccordement ne présente pas de discontinuités. Des traitements complexes à base de fenêtres de Hanning et de transformées de Fourier entrent en jeu.

► *Prosodie* : Lors de la lecture d'un texte, la voix subit des variations à l'échelle de la syllabe, du mot, du groupe de mots, de la phrase, du paragraphe. Ces variations, volontaires ou non, contribuent au naturel, à l'intelligibilité et à l'expressivité du discours. Elles portent sur les trois principaux paramètres de variation de la parole : hauteur, intensité, durée, auxquels il conviendrait d'ajouter le timbre, beaucoup plus difficile à modéliser (cf. Lacheret-Dujour & Beaugendre 1999, Lacheret-Dujour 2002). En l'absence de prosodie, la voix paraît plate et monotone. Le modèle actuel ne traite pour l'instant que la déclinaison et les accents démarcatifs.

► *Générateur de parole* : Le module générateur de parole convertit la chaîne phonémique en un signal de parole par concaténation des diphones de la base. Les variations prosodiques calculées sont appliquées en temps réel. Le signal de parole est envoyé en continu vers les haut-parleurs à fréquence d'échantillonnage constante.

Le synthétiseur Kali présente un potentiel d'évolution important. Grâce à leurs règles déclaratives structurées, l'analyse syntaxique et la transcription graphème-phonème sont améliorées progressivement au cours de l'étude de corpus. La parole de base peut encore gagner en naturel et en agrément d'écoute par le fignage des bases de diphones et des programmes de traitement. La prosodie devrait bénéficier d'une amélioration spectaculaire avec l'utilisation de patrons prosodiques en adéquation avec la nature et l'enchaînement des groupes intonosyntaxiques, puis avec la prise en compte d'indices discursifs (focus, attitudes, etc.). La synthèse de l'anglais, en cours de développement, va prochainement être ajoutée. Cet objectif présente de nombreux avantages tant pour les applications industrielles que pour la recherche.

1.5. Le domaine de recherche central du CRISCO se décline en termes de strates syntaxiques⁴. La connexion interpropositionnelle et sa « face cachée », le détachement, viennent en premier (cf. ci-dessous §2). Le verbe est abordé en français, allemand, finnois et en basque à différents points de vue : syntaxique, sémantique, morphologique et en fonction des différentes opérations prédicatives qui lui sont applicables (§3). Les classes d'adjectifs du français sont abordées en corrélation avec celles des adverbes et la réduction de leur polysémie en contexte nominal est étudiée à l'aide du *D.E.S.* couplé avec l'exploitation de la composante catégorisée de la base de données textuelles FRANTEXT de l'ATILF (Universités de Nancy 1 et 2) (§4). La corrélation entre la syntaxe de la phrase circonstanciée et son modelé prosodique est étudiée en termes de phonétique acoustique, de synthèse vocale et d'imagerie cérébrale (§5). La significativité des unités linguistiques est également un domaine abordé au CRISCO de manière diversifiée (§6). Enfin, compte tenu de l'orientation « sémantique » dominante de l'équipe, le CRISCO est un espace privilégié d'observation des controverses épistémologiques actuelles sur l'autonomie de la grammaire (point de vue « formaliste ») ou sa dépendance de la fonction de communication (point de vue « fonctionnaliste ») ou de celle de conceptualisation (point de vue de la grammaire cognitive, §7).

2. La connexion inter- et intrapropositionnelle et le détachement

2.1. Le CRISCO consacre une étude approfondie à un certain nombre de marqueurs grammaticaux réunis sous l'étiquette générale de « connecteurs ». Le terme est pris dans un sens large et englobe tous les marqueurs qui assurent une connexion syntaxique, soit au niveau inter-propositionnel (conjonctions de subordination), soit au niveau inter-phrastique (certains adverbes, notamment ceux que l'on englobe habituellement dans la catégorie des connecteurs pragmatiques), soit au niveau inter-syntagmatique (prépositions). Les langues concernées sont le français (y compris dans sa dimension diachronique), l'anglais et l'allemand.

La démarche adoptée est de type sémasiologique et vise à rendre compte des valeurs attestées des marqueurs à partir de deux critères essentiels : une valeur invariante attachée au marqueur et le jeu des indices contextuels qui permettent de générer le sens à partir de cet invariant. Ces principes de base sont partagés par l'ensemble des membres de cette opération même si ces derniers peuvent s'inscrire dans des cadres théoriques divergents : théorie des opérations énonciatives, systématique énonciative, linguistique fonctionnelle. A titre d'exemple, dans le cadre de la théorie des opérations énonciatives, cet invariant de base sera défini comme une forme schématique suffisamment abstraite pour pouvoir générer l'ensemble des interprétations possibles. Cette forme schématique est renseignée par les divers éléments contextuels et aboutit à un résultat sémantique. Dans le cas des prépositions par exemple, aucune valeur, pas même la valeur spatiale, n'est considérée comme première ; toutes sont au contraire traitées comme construites en discours à partir de leur invariant de base et de l'agencement linguistique dans lequel elles apparaissent. Les différentes formes schématiques définies à partir d'un certain nombre d'opérations centrales telles que le repérage et ses différentes valeurs (identification, différenciation, disjonction) en association avec les concepts de prédication d'existence et de prédication de propriété (respectivement

⁴ Le CRISCO aborde également des questions de terminologie linguistique (cf. Neveu 2000), de variation de l'oral (cf. Bougy 2000, Camugli Gallardo 2002), de typologie des textes (cf. Legallois coord. 2003), et de traductologie (cf. Durieux 2002). En outre, dans une perspective typologique, le CRISCO a mis en place un projet d'une base de données en typologie des langues (http://www.unicaen.fr/typo_langues) dont le principal atout sera la recherche et le croisement de propriétés typologiques (Richard Renault).

dénommés quantification et qualification dans le modèle culiolien) sont informées par les propriétés notionnelles et situationnelles des termes repère et repéré et déclenchent telle ou telle interprétation, par exemple de *in*, *on*, *over*, *above*, etc. Les recherches menées actuellement au sein de cette opération sont orientées plus spécifiquement vers un type de repérage particulier, la localisation.

Les résultats de ces travaux ont fait l'objet de plusieurs publications individuelles (cf. Gilbert 2004a, 2004b, Le Querler 2003). De surcroît, on en trouvera une bonne illustration dans le premier numéro de la revue *Syntaxe & Syntaxique* (Guimier, coord. 2000), recueil d'articles émanant de notre groupe. Les principes énoncés ci-dessus sont largement appliqués à un certain nombre de marqueurs du français (*où*, *tandis que* et *alors que*, *or*, *lors* et l'apparition de *lorsque* dans l'histoire du français), de l'anglais (*for*, *as*, *why*) et de l'allemand (prépositions, adverbes et particules verbales). Ce recueil présente un ensemble de contributions qui, tout en s'inscrivant dans des cadres théoriques différents, ont un même objectif : derrière la diversité des emplois attestés, retrouver l'unité du marqueur et décrire la genèse de ses diverses interprétations.

2.2. L'objectif du volet consacré à la discontinuité syntaxique en français (cf. Neveu 1998, coord. 2000, 2002, coord. 2003) est de travailler à un état des lieux de la notion de détachement, et à une mise au point sur les principaux événements linguistiques susceptibles d'entrer dans son champ d'application (notamment les constructions appositives, disloquées, topicalisées, clivées, incisives, vocatives). Cinq problématiques sont développées dans ce cadre : (i) la question des choix terminologiques (*détachement*, *adjonction*, *disjonction*, *discontinuité*, *incidence*, *séparation*, etc.), et celle de leurs implications méthodologiques ; (ii) la question du principe structurel présumé commun aux divers types de segments détachés qui justifierait leur réunion dans une même problématique syntaxique ; (iii) la question de la connexité des segments détachés avec leur environnement linguistique (connexité morphosyntaxique et dépendance réactionnelle du segment avec le reste de la séquence *vs* connexité sémantico-pragmatique et relation de juxtaposition et d'interdépendance entre les constituants) ; (iv) la question de la position des segments dans l'énoncé et celle de la nature des opérations linguistiques associées à telle ou telle structure détachée (modalisation, prédication seconde, topicalisation, thématisation, rhématisation, etc.) ; (v) la question du détachement comme dispositif syntaxique de la réflexivité linguistique (gloses), et comme figure de l'hétérogène (faits de stratification énonciative, tâtonnements syntagmatiques, ruptures, ratages, etc.).

3. Le verbe : syntaxe, sémantique, morphologie et opérations prédicatives

3.1. La prédication est fondamentalement un acte de discours qui en français est pris en charge habituellement par un verbe (prédicatif) et plus rarement par un adjectif ou un nom en emploi prédicatif. Les théories grammaticales le mieux outillées pour rendre compte de l'exercice de l'acte de prédication verbale dans toutes ses dimensions sont les modèles formels d'inspiration fonctionnelle, dont la *Functional Grammar* élaborée par Simon Dik et la *Role and Reference Grammar* développée par Robert van Valin sont actuellement les illustrations les plus vivantes .

A la lumière de ces deux cadres théoriques, J. François (2002, 2003) met en évidence les propriétés aspectuelles et participatives de la « prédication verbale » dans le sens particulier que confère à ce terme la *Functional Grammar*, c'est-à-dire la strate de la structure sous-jacente de la phrase constituée du prédicat verbal et de son cadre prédicatif, domaine de rattachement d'opérateurs et de satellites spécifiques (en particulier la segmentation d'un

procès en phases et son ancrage spatio-temporel). Ces deux types de propriétés constituent des facteurs sémantiques interagissants de la transitivité syntaxique.

La mise en évidence de ces propriétés permet l'élaboration d'un classement hiérarchisé des prédications verbales dont quatre propriétés primaires, les trois propriétés participatives de relationalité, agentivité et causativité et la propriété aspectuelle de transitionnalité, constituent la charpente. En outre l'analyse de la perte de prédictivité de différentes classes d'entrées verbales conduit à l'élaboration d'un continuum de prédictivité vs. auxiliaire fondé sur les relations de réction mutuelle qu'entretiennent deux à deux les principaux représentants de ces classes d'entrées verbales non prédictives.

Enfin à travers trois études contrastives entre le français et l'allemand, l'auteur montre que c'est à partir de la comparaison avec une autre langue fondée sur le recours à un large corpus de traductions que les traits caractéristiques du lexique verbal français se dégagent le plus clairement. A travers ces études détaillées, l'objectif ultime de cette recherche est d'opérer un double rapprochement, d'une part entre les entreprises de structuration du lexique verbal centrées soit sur les propriétés, soit sur les propriétés participatives, d'autre part entre les approches formelles caractérisées par le primat de la syntaxe et les approches fonctionnelles qui conçoivent la syntaxe comme un outil de transmission du sens des expressions linguistiques.

D'autres études menées au CRISCO ont une thématique apparentée, qu'il s'agisse du spectre des emplois des morphèmes polycatégoriels de l'allemand, tels *aus*, *bei*, *mit* ou *von* susceptibles d'assumer une fonction de préverbe, de préposition ou de substitut (Krause, 2002) ou de la question de la catégorie des participes du finnois (emplois adjectivaux, verbaux et nominaux) et du français (emplois adjectivaux et verbaux), abordée dans le cadre des catégories fonctionnelles du modèle minimaliste de la grammaire générative (Renault 2004a, 2004b).

3.2. Plusieurs chercheurs du CRISCO sont engagés en collaboration avec des collègues de deux autres unités membres de l'*Institut de Linguistique Française*, le LATTICE (Paris) et l'ERSS (Toulouse) dans l'étude des relations entre sens et constructions verbales de verbes polysémiques dans le cadre théorique des « grammaires de construction » (cf. Goldberg 1995) et de la construction dynamique du sens (cf. Victorri & Fuchs 1996), à partir de l'analyse linguistique de corpus et l'utilisation d'outils informatiques d'analyse syntaxique, de représentation géométrique du sens (à l'aide du *D.E.S.*) et de calcul du sens d'un verbe en contexte. Actuellement, les recherches menées sur des verbes polysémiques tels que *jouer*, *passer* et *participer* portent tout d'abord sur l'analyse linguistique de la corrélation entre la polysémie et les cadres distributionnels liés à ces unités verbales, en se référant à la base de données textuelles Frantext et à des corpus plus spécialisés. D'autre part, la passage à la représentation géométrique du sens avec ajout automatique du contexte et de la configuration actancielle se fera par le biais d'analyseurs syntaxiques dont nous disposons déjà, et par la construction d'une interface qui exploitera le *D.E.S.* pour circonscrire l'ensemble des requêtes à formuler vers les corpus, puis pour construire les représentations du sens en tirant parti des données paradigmatiques du dictionnaire, et des données syntagmatiques issues des traitements d'analyse syntaxique.

4. L'adjectif et l'adverbe

4.1. La syntaxe et la sémantique des adjectifs et de leurs dérivés adverbiaux ont fait l'objet au CRISCO d'analyses pointues dans le cadre de la préparation en 2001 du colloque du CRISCO *L'adjectif en français et à travers les langues* (François, coord., sous presse).

Certains adjectifs, lorsqu'ils sont antéposés, perdent leur valeur qualificative, et n'autorisent alors (du moins avec certains substantifs) aucune relation de paraphrase entre leur position prénominale et leur position postnominale. Ainsi, *une ancienne armoire* n'est pas *une armoire ancienne*, *un simple professeur* n'est pas *un professeur simple*, *un faux prêtre* n'est pas *un prêtre faux*, etc. (cf. Salles, sous presse). La question intéressante est alors celle de l'homonymie : a-t-on affaire au même adjectif ou à deux adjectifs différents ? J.C. Milner considère que ces adjectifs ne sont pas à place variable, comme le pensaient les grammairiens de la première moitié du siècle, mais des adjectifs à place fixe qu'il qualifie de « modaux ». Cependant l'existence de cas intermédiaires suggère l'existence d'un **continuum** entre des emplois modaux et des emplois qualificatifs.

Les adjectifs « primaires » prennent fréquemment une valeur intensive en position d'épithète antéposée (*de gros avantages*, *de larges bénéfices*, *de maigres revenus*, *une faible amélioration*, *une nette progression*, *un léger retard*, ...). Cette valeur intensive leur confère un rôle particulier dans l'enchaînement argumentatif dont la description impose de faire intervenir la dimension axiologique de l'argumentation (cf. Lenepveu, Salles & François, 2002).

Un type très particulier d'adjectif est représenté par les « épithètes de nature », par exemple *le bouillant Achille*, *une méchante sorcière*, *l'épaisse forêt*, qui se prêtent à une étude linguistique et non seulement rhétorique (cf. Delente, sous presse). Ces épithètes révèlent des stratégies discursives spécifiques, mises en œuvre dans les discours poétiques, épidiotiques, ainsi que dans les contes merveilleux pour enfant. Elles se laissent définir par deux caractéristiques : (1) la nécessaire antéposition, (2) la relation analytique ou présentée comme telle entre l'épithète de nature et le substantif. Dans *un petit nain*, elle explicite une propriété analytique du N ; dans *un dangereux terroriste*, *un habile magicien* et *l'épaisse forêt*, elle présente comme analytique une propriété stéréotypique. Le syntagme nominal comprenant une épithète de nature constitue une séquence qui est l'objet d'un certain degré de figement aux plans prosodique, syntaxique et sémantique. Alors que l'adjectif postposé déclenche une opération consistant à décrire un objet, l'épithète de nature déclenche, pour son interprétation, le mécanisme cognitif qu'est *l'évocation*.

Quant aux relations syntaxiques et sémantiques entre adjectifs et adverbes en *-ment* morphologiquement dérivés de ces adjectifs, V. Lenepveu (2002) montre que la corrélation adjectif qualificatif ~ adverbe de manière ne constitue qu'un cas particulier d'une corrélation plus générale entre deux types de fonctionnement des adjectifs et deux types de fonctionnement des adverbes en *-ment* dérivés de ces adjectifs. La corrélation est sémantique, mais aussi positionnelle : à l'adjectif qualificatif correspond un fonctionnement d'adverbe de manière tandis qu'à l'adjectif antéposé qui peut prendre des valeurs diverses, correspondent différents fonctionnements de l'adverbe distinct de celui d'adverbe de manière (*Il a formulé une réponse simple / Il a répondu simplement*, *Elle portait une simple robe / Elle portait (tout) simplement une robe*).

4.2. Le CRISCO a engagé en 2000 dans le cadre de l'action « *Cognitive* » du Ministère de l'Enseignement et de la Recherche un projet dédié à l'étude de l'espace sémantique des adjectifs et de leur réduction en contexte nominal, en collaboration avec le laboratoire *LATTICE* de Paris et le laboratoire de psychologie cognitive *LaCo* de Poitiers. Le but de ce projet était d'une part de calculer la déformation de l'espace sémantique des adjectifs polysémiques en contexte de rection nominale, et d'autre part d'associer cette étude à l'évaluation de l'accessibilité et du poids des relations synonymiques adjectivales hors contexte et en contexte nominal chez des sujets francophones natifs d'âge et de formation différents. La partie théorique exploratoire a permis de bâtir différents modèles

mathématiques d'élargissement de la représentation du sens hors contexte à une représentation du sens de l'adjectif en contexte nominal, à partir des relations existant dans le *D.E.S.* et des données de cooccurrences récupérées de manière automatique dans le corpus Frantext catégorisé et dépouillées par un programme construit pour les besoins de cette étude. La partie psycholinguistique de l'étude a mis en évidence, pour la partie hors contexte, les différentes étapes de l'organisation de l'environnement synonymique chez les sujets en fonction de leur âge, et que l'on peut résumer ainsi : lors du passage du primaire au secondaire, c'est l'enrichissement du lexique qui agit sur l'organisation de cet environnement ; vers la fin du secondaire, on mesure les effets de l'enrichissement sémantique des mots, et enfin, par la suite, on peut observer chez certains sujets l'acquisition d'une expertise qui complète l'acquis précédent (cf. François, Victorri & Manguin, sous presse ; Manguin, François & Victorri, sous presse ; François, Manguin, Cordier & Ross à paraître).

5. Intonosyntaxe et pragmatique

Le problème de la correspondance entre les aspects fonctionnels et formels du langage pour décrire et modéliser la **prosodie de la parole** constitue l'une des thématiques centrales de la phonologie cognitive aujourd'hui, c'est autour d'elle que s'articulent les travaux du CRISCO dans le prolongement de Lacheret & Beangendre (1999).

La prosodie se manifeste principalement, dans le signal de parole, par trois paramètres : la fréquence fondamentale (estimation du son laryngien à partir du signal acoustique à un instant *T*), la durée (mesure d'un intervalle de temps nécessaire pour émettre le signal de parole et les unités qui le constituent) et l'intensité (l'énergie contenue dans le signal). Au niveau abstrait – ou phonologique – le terme *prosodie* désigne une structure articulée autour de deux mécanismes imbriqués et à portée variable selon les langues : l'accentuation et l'intonation. En d'autres termes, alors que les objets phonétiques désignent des paramètres physiques concrets, le niveau de traitement phonologique correspond à l'analyse du système prosodique, variable d'une langue à une autre, celui là même qui donne lieu à l'apprentissage d'un jeu de règles par l'enfants en phase d'acquisition du langage.

En pratique donc, se pose le problème de l'interface entre ce niveau de traitement et les niveaux syntaxique, sémantique et pragmatique: il s'agit de montrer : (i) dans quelle mesure la prosodie appartient à la grammaire, au même titre que la syntaxe et la sémantique (ii) de proposer des représentations formelles susceptibles d'en rendre compte, (iii) de poser un ensemble d'arguments qui militent en faveur de la réalité psychologique de ces représentations, (iv) d'en proposer une interprétation pragmatique qui rende compte aussi rigoureusement que possible de la **fonction communicative** centrale de la prosodie. Cette fonction se décline en trois points centraux : une **fonction structurale** : la prosodie agit comme un dispositif organisateur de mise en forme du matériau langagier et de mise en perspective de l'information ; une fonction de contextualisation, *i.e.* d'intégration des messages à leur contexte ; une fonction énonciative : la prosodie porte les traces des diverses façons dont les locuteurs s'investissent personnellement dans leurs messages (cf. Lacheret-Dujour, 2002).

Si cette thématique est par essence dévolue à la linguistique et à la phonétique, elle s'inscrit également dans une perspective transdisciplinaire qui conduit naturellement à la mise en place de collaborations externes à notre laboratoire avec des disciplines connexes. Ainsi, la psychologie expérimentale et les neurosciences permettent de vérifier l'adéquation des hypothèses théoriques défendues avec la réalité cognitive, *i.e.* les bases neurales qui sous-tendent la production et la perception de la prosodie dans les langues. Dans ce contexte, A.

Lacheret-Dujour poursuit deux collaborations, l'une avec le laboratoire de neurosciences cognitives de Marseille sur le marquage prosodique du focus contrastif en français en utilisant les méthodes comportementale et électrophysiologique (PEs), l'autre, en imagerie cérébrale, avec le centre CYCERON de Caen sur la prosodie affective. Le système de synthèse vocale KALI et l'architecture environnante sont utilisés en amont et en aval de ces travaux pour l'enregistrement des stimuli et la validation des hypothèses formulées dans une tâche de génération automatique de la parole (cf. Morel & Lacheret-Dujour 1998, 2001).

6. Unités linguistiques et significativité

Si les travaux du CRISCO portent globalement sur des unités explicitement significatives (des connecteurs aux unités macrosyntaxiques) ou non significatives (par exemple les diphtonges mis en œuvre dans la synthèse vocale de KALI), la thème de la significativité des unités linguistiques y est toutefois aussi abordé dans deux perspectives très différentes.

Première question : les unités submorphémiques peuvent-elles être dotées d'un signifié, contrairement à la définition classique du morphème comme unité significative minimale ? La théorie des cognèmes mise en place par D. Bottineau (2002, ; à paraître) propose une modélisation de l'interaction cognitive locuteur / allocutaire par l'acte de langage en s'appuyant sur l'analyse de la structure morphologique des grammèmes dans les langues naturelles. Certaines d'entre elles, de types et origines divers (romanes, germaniques, altaïques, atlantiques, basque), possèdent des grammèmes formés d'une combinaison d'éléments submorphémiques minimaux (phonème et/ou du graphème) : au sein d'un grammème, un phonème, asémantique par principe dans le lexique, devient un cognophone, marqueur d'un invariant sémantique, le cognème (comme *th-* à valeur de reprise en anglais dans *the, this* etc.). La sémantique des cognèmes n'est ni référencielle ni phonosymbolique mais opératoire : à chaque marqueur correspond une procédure cognitive (*i* pour jonction, *a* pour disjonction, *s* pour présent, *t* pour passé, etc) pouvant entrer dans la composition d'un opérateur (*this = th + i + s* : jonction de la reprise anaphorique d'un référent repéré et de la sélection présente d'une notion lexicale pour le nommer). Cette stratégie permet à un locuteur d'amorcer la procédure interprétative chez l'allocutaire en inscrivant dans la structure même du grammème la nature des opérations de mise en relation à instaurer entre les notions lexicales en vue d'assembler correctement les unités sémantiques et d'en repérer les référents (détermination, prédication, actance, aspect, temps, modalité).

Deuxième question : où l'étude frénétique des anagrammes a-t-elle conduit F. de Saussure ? F. Gandon, qui poursuit une vaste étude sémiotique des *Cahiers d'anagrammes* de Saussure consacrés à Lucrèce, du triple point de vue de leur relation aux disciplines indo-européennes, au cours de linguistique générale, aux recherches mythographiques présente une « machinerie anagrammatique [qui] fait craquer les cosses rétives et les anciens carcans des mots ; elle prépare un monde lisible : purifié — enfin — du sens » (2002 :202). Saussure semble avoir vécu « une illumination noire, au service du non sens » (p.27). L'expérience des anagrammes lui a fait « atteindre un prédicat aussi proche que possible de zéro, aux plans à la fois du signifié (absence de pertinence du contenu second) et du signifiant (annulation terme à terme par le principe de parité) » (ibid.).

Fascinant contraste : D'un côté un mot peut entrer, par les cognèmes qu'il véhicule subrepticement, dans un schème opératoire qui subvertit les oppositions supposées minimales, de l'autre les anagrammes livrent un objet a priori linguistique, mais exempté de tout symbolisme, miroir brisé juste propre à refléter des éclats de sens.

7. Un observatoire des nouvelles grammaires fonctionnelles et cognitives

Finalement, le CRISCO se conçoit aussi comme un observatoire du débat épistémologique en cours entre, à un premier niveau, les «*formalistes*» et les «*fonctionnalistes*» (cf. Darnell, coord. 1999), c'est-à-dire respectivement les représentants d'une conception des expressions linguistiques soit déconnectée de leur usage en contexte (la «*grammaire interne*» de Chomsky) soit inversement «*fondée sur l'usage*», et à un second niveau, parmi ceux-ci, les tenants d'un fonctionnalisme prioritairement orienté vers la fonction de communication langagière ou vers celle d'expression de l'activité de cognition.

Un rapprochement spectaculaire s'est effectué dans les deux dernières décennies du 20^e siècle entre les orientations fonctionnelle et cognitive en théorie grammaticale. Ce rapprochement résulte du développement de la typologie fonctionnelle des langues qui ne tient plus seulement compte de motivations internes au langage, mais a développé l'analyse du jeu complexe de compétitions et de coopérations entre motivations internes et externes (d'ordre cognitif, énonciatif et social) qui stabilise les langues ou les fait évoluer. Les analyses de R. Langacker sur la transitivité ou la catégorisation syntaxique (1991) vont dans le même sens que celles de Givón (1995) ou Croft (1990) parce que ces auteurs intègrent le point de vue des représentations cognitives dans leur approche fonctionnelle. Des linguistes comme R. Langacker, T. Givón ou W. Croft partagent globalement une perspective «*constructionniste*» intégrant représentations cognitives et approche fonctionnelle. Cette intégration permet de comprendre pourquoi les langues, tout en stabilisant des structures arbitraires, conservent un fond d'iconicité.

Au delà, l'entreprise de repérage des fondations du langage de R. Jackendoff (cf. Jackendoff 2002) favorise un rapprochement plus vaste entre les théories de la grammaire «*fondées sur l'usage*» (le regroupement des linguistiques fonctionnelles et cognitives selon M. Tomasello) et le modèle d'«*architecture parallèle*», c'est-à-dire de génération de structures de trois types, phonologiques, syntaxiques et sémantiques-conceptuelles qu'il préconise lui-même (cf. François 2004).

En France, diverses théories plus ou moins récentes (psychomécanique du langage de G. Guillaume, théorie des opérations énonciatives d'A. Culioli, grammaire méta-opérationnelle d'H. Adamczewski, théorie de l'interlocution de C. Douay et D. Roulland entre autres) relèvent à divers titres de modélisations cognitives du langage, des langues et de la communication, sans nécessairement revendiquer leurs affinités avec cette orientation de recherche et en se démarquant notablement des linguistiques cognitives reconnues sous ce label. La confrontation de ces différentes approches permet d'étudier l'état actuel des relations entre linguistique et sciences cognitives ainsi que l'émergence de nouvelles problématiques fédérant leurs orientations, compatibles ou divergentes selon les cas. Au cœur de ces controverses on retrouve fondamentalement le débat sur la relation entre la (morpho-)syntaxe et la (pragma-)sémantique : simple **correspondance** (*linking*, point de vue de Jackendoff) ou **indissociabilité** dans une vaste «*sémantaxe*» (point de vue des fonctionnalistes et de la grammaire cognitive) . Il semble actuellement que les approches intermédiaires que Croft (1995) qualifie de «*fonctionnelles et formelles*» et Butler (2003) de «*structurales fonctionnelles*» soient les mieux adaptées aux problématiques abordées au CRISCO (cf. François, coord. 2003).

Références

- BOTTINEAU, Didier (2002), « Les cognèmes de l'anglais et autres langues », in : A. Ouattara (ed), Actes du Colloque International « *Parcours énonciatifs et parcours interprétatifs* », Université de Tromsø (Norvège), 26-29 octobre 2000, Ophrys.
- BOTTINEAU, Didier (à paraître), « Périphrases verbales et genèse de la prédication en langue anglaise ». In : H. Bat Zeev Schyldkrot & N. Le Querler (coord.) *Les périphrases verbales*. Editeur non spécifié.
- BOUGY, Catherine (2000), « Le langage ordinaire d'Italiens implantés en Basse-Normandie », *Annales de Normandie* n° 29, Caen, Musée de Normandie, p. 271-286.
- BUTLER, Christopher (2003), *Structure and function. A guide to three major structural-functional theories*. Amsterdam: Benjamins. [Part 1: *Approaches to the simplex clause*; Part 2 : *From clause to discourse and beyond*]
- CAMUGLI GALLARDO (2002), « La langue des jeunes en Italie — Guide à une étude de terrain ». *Cahier du CRISCO 7*. CRISCO-CNRS, Université de Caen.
- CORDIER, Françoise & FRANÇOIS, Jacques (coord., 2002), *Catégorisation et langage*. Paris : Hermès. Collection de sciences cognitives, Paris : Hermès.
- CORDIER, Françoise, VICTORRI, Bernard & FRANÇOIS, Jacques (coord., 2001), *Sémantique du lexique verbal. Syntaxe & Sémantique 2*. Caen : Presses Universitaires de Caen.
- CROFT, William (1990), *Typology and universals*. Cambridge : Cambridge University Press.
- CROFT, William (1995), «Autonomy and functional linguistics ». *Language* 71, 490-532.
- DARNELL , Michael *et al.* (coord. 1999), *Functionalism and formalism in linguistics*. Volume 1 : *General papers*. Amsterdam: Benjamins
- DELENTE, Eliane (sous presse), L'épithète de nature ou « Les terroristes sont-ils dangereux? ». In. J. François (coord.).
- DURIEUX, Christine (2002), « Le traitement du figement lexical en traduction ». *Cahiers de lexicologie* 82. Champion, Paris. 82,193-207
- DUTOIT D, NUGUES P., P. DE TORCY (2003) « The Integral Dictionary : a lexical network based on computational semantics », *Springer Ed., May 2003, ICCSA International Conference on Computational Science and its Applications*, Calgary, Canada
- DUTOIT, Dominique & POIBEAU, Thierry (à paraître) : « Inferring Knowledge from a Large Semantic Network », Actes de *Conference on Computational linguistics, COLING, TAIWAN*, août 2002.
- FRANÇOIS, Jacques (2002), « La catégorisation verbale : morphologie, syntaxe, sémantique ». In F. Cordier & J. François (coord.) Chapitre 2, pp. 53-82.
- FRANÇOIS, Jacques (2003), *La prédication verbale et les cadres prédicatifs*. Louvain: Peeters, [Bibliothèque de l'Information Grammaticale 54].
- FRANÇOIS, Jacques (2004) « Le fonctionnalisme linguistique et les enjeux cognitifs ». In C. Fuchs (ed.) *La linguistique cognitive*. Editions de la Maison des Sciences de l'Homme & Ophrys.
- FRANÇOIS, Jacques (coord., 2003), *Aspects de la Role and Reference Grammar. Cahier du CRISCO 14*. CRISCO-CNRS, Université de Caen.
- FRANÇOIS, Jacques (coord., sous presse), *L'adjectif en français et à travers les langues*. Caen : Presses Universitaires de Caen.
- FRANÇOIS, Jacques, Manguin, Jean-Luc & Victorri, Bernard (2003), « La réduction de la polysémie adjectivale en cotexte nominal : une méthode de sémantique calculatoire ». *Cahier du CRISCO 13*. CRISCO-CNRS, Université de Caen.
- FRANÇOIS, Jacques, VICTORRI, Bernard & MANGUIN, Jean-Luc (sous presse), « Polysémie adjectivale et synonymie : L'éventail des sens de *curieux* étudié à l'aide de

- méthodes informatiques combinant un dictionnaire électronique de synonymes et des extraits de corpus ». In O. Soutet (éd.), *La polysémie*, Actes du colloque de Paris IV, novembre 2000. Presses de l'Université de Paris-Sorbonne.
- FRANÇOIS, Jacques, MANGUIN, Jean-Luc, CORDIER, François & ROSS, Christine (à paraître) : « Le dictionnaire électronique des synonymes du CRISCO – Méthode d'utilisation et exploitation psycholinguistique », in R. Canac-Marquis (éd.), *Actes du colloque Multimédia, Internet et Etudes françaises II*, Vancouver, mars 2003. Editeur non spécifié.
- GANDON, Francis (2002), *De dangereux édifices. Saussure lecteur de Lucrèce. Les cahiers d'anagrammes consacrés au De Rerum natura*, Peeters [Bibliothèque de l'information grammaticale].
- GILBERT, Eric (2004a), « *Across, by et through*. Considérations sur les conditions de représentation métalinguistique des prépositions », *Anglophonia*, 14.
- GILBERT, Eric (2004b), « Ebauche d'une formalisation des prépositions *in, on et at* », *Cycnos* 21(1).
- GIVÓN, Talmy (1995), *Functionalism and grammar*. Amsterdam : Benjamins.
- GOLDBERG, Adele (1995), *Constructions : A Construction Grammar approach to argument structure*. Chicago: Chicago University Press.
- GUIMIER, Claude (coord., 2000), *Connecteurs et marqueurs de connexions, Syntaxe & Sémantique 1*, Presses Universitaires de Caen.
- GUIMIER, Claude, 2002, La conjonction complexe *as if*. In N. Le Querler et G. Kleiber, (coord.) *Traits d'union, Actes des 12èmes rencontres linguistiques en Pays Rhénan*, Presses Universitaires de Caen, pp. 105-119.
- JACKENDOFF, Ray (2002), *Foundations of language — Brain, Meaning, Grammar, Evolution*. Oxford, New-York : Oxford University Press.
- KRAUSE, Maxi, 2000, Vers une grammaire des prépositions, substituts et particules verbales de l'allemand : bilan et perspectives, *Syntaxe & Sémantique, 1, Connecteurs et marqueurs de connexions*, C. Guimier, éd., PU Caen, pp. 213-242.
- LACHERET-DUJOUR, Anne (2002), *Modélisation prosodique du français parlé : analyse de la substance, représentation formelle, interprétation symbolique*, Dossier d'habilitation à diriger des recherches, Université de Paris X, Nanterre, juillet 2002.
- LACHERET-DUJOUR Anne, BEAUGENDRE Frédéric (1999), *La prosodie du français*, Paris : CNRS Editions.
- LANGACKER, Ronald (1991), *Foundations of Cognitive Grammar, Volume 2 : Descriptive applications*. Stanford : Stanford University Press.
- LEGALLOIS, Dominique (coord. 2003), *Unité(s) du texte* (Journée scientifique du CRISCO, 6 décembre 2002.) *Cahier du CRISCO 12*. CRISCO-CNRS, Université de Caen.
- LEGALLOIS, Dominique (sous presse) : « Synesthésie adjectivale, sémantique et psychologie de la forme : la transposition au cœur du lexique ». In J. François (coord.)
- LENÉPVEU, Véronique (2002), « Adjectifs et adverbes: une corrélation syntactico-sémantique », *Le français moderne*, 70, 1, 45-70.
- LENÉPVEU, Véronique, Salles, Mathilde & François, Jacques (2002), « La catégorisation adjectivale et adverbiale ». In Cordier & François (coord.) Chapitre 3, pp. 83-105
- LE QUERLER, Nicole, 2000, Le connecteur *où*, *Syntaxe & Sémantique, 1, Connecteurs et marqueurs de connexions*, C. Guimier, éd., PU Caen, pp.113-132
- LE QUERLER, Nicole, 2003, « Actance verbale et actance nominale : le cas des agents en *par* ou en *de* », *Travaux Linguistiques du CERLICO*, 16.
- MANGUIN, Jean-Luc, FRANÇOIS, Jacques & VICTORRI, Bernard (sous presse), « Polysémie adjectivale et rection nominale : quand *gros* et *gras* sont synonymes » in J.

- François (coord.), *L'adjectif en français et à travers les langues*. Presses Universitaires de Caen.
- MOREL, Michel & LACHERET-DUJOUR, Anne (1998), « Utilisation d'une structure arborescente pour une hiérarchisation fine des règles de transcription graphème-phonème », *Actes des XXIIèmes Journées d'Etudes sur la Parole*, 15-19 juin 1998, Martigny, Suisse, pp. 151-154.
- MOREL, Michel, LACHERET-DUJOUR, Anne (2001) « Le logiciel de synthèse vocale Kali : de la conception à la mise en œuvre », *Traitement Automatique des Langues*, Ch. D'Alessandro (dir.), Paris: Hermès, vol. 42, pp. 193-221.
- NEVEU, Franck (1998) *Études sur l'apposition*, Paris, Honoré Champion, coll. « Grammaire et linguistique », 286 pages.
- NEVEU, Franck (2000), *Lexique des notions linguistiques*. Paris, Nathan.
- NEVEU, Franck (dir. 2000), « Nouvelles recherches sur l'apposition » *Langue Française*, n° 125, Paris, Larousse, 128 pages.
- NEVEU, Franck (2002), « L'ajout et la problématique appositive — Détachement, espace phrastique, contextualité », in J. Authier-Revuz & M.-C. Lala (éd), *Figures d'ajout : phrase, texte, écriture* Paris, Presses de La Sorbonne nouvelle : 111-122.
- NEVEU, Franck (2003), « La glose et le système appositif », in A. Steuckardt & A. Niklas-Salminen, *Le Mot et sa glose*, Aix-en-Provence, Presses de l'Université de Provence, *Langues et langages*, 9 : 143-167.
- NEVEU, Franck (dir. 2004) « Linguistique du détachement », *Cahiers de praxématique* n°40, Montpellier, Université Montpellier III.
- RENAULT, Richard (2004a), « Morphosyntaxe des participes finnois », *Études Finno-ougriennes* n°36, ADEFO / l'Harmattan, Paris.
- RENAULT, Richard (2004b), « La catégorie des participes en finnois et en français », *Actes du 6^e colloque contrastive français-finnois*, Université de Helsinki.
- SALLES, Mathilde (2001), Hypothèse d'un continuum entre les adjectifs 'modaux' et les adjectifs qualificatifs, *L'Information grammaticale*, 88, pp. 23-27.
- SALLES, Mathilde (sous presse), « Adjectifs 'modaux' et adjectifs 'qualificatifs' », In J. François (coord.).
- VICTORRI, Bernard & FUCHS, Catherine (1996), *La polysémie, construction dynamique du sens*. Paris : Hermès.