



Le projet d'étude sur l'évolution des déterminants ELICO

Lucia M. Tovenà et Jacques Jayez

Pour citer cet article

Tovenà Lucia M. et Jacques Jayez (2010). Le projet d'étude sur l'évolution des déterminants ELICO. Dans Lucia M. Tovenà (éd.) *Déterminants en diachronie et synchronie*. Paris : Projet ELICO Publications, p. 1-13.

<http://elico.linguist.univ-paris-diderot.fr/livre-elico.html>

LE PROJET D'ETUDE SUR L'EVOLUTION DES DETERMINANTS ELICO¹

L.M.Tovena et J. Jayez

tovena@linguist.jussieu.fr

jjayez@ens-lsh.fr

Résumé

Dans ce texte, nous présentons brièvement le projet *Evolution linguistique et corpus* (ELICO), qui a servi de cadre général aux recherches proposées dans ce volume. Le corpus réalisé consiste en une collection d'occurrences de déterminants présentés dans leur environnement, lequel a été annoté avec un certain nombre de propriétés. La collection couvre six siècles (13^{ème} – 18^{ème}) et se compose de textes de différents genres.

Dans ce texte, nous présentons d'abord brièvement le projet *Evolution linguistique et corpus* (ELICO), qui a servi de cadre général aux recherches proposées dans ce volume, avant de résumer les points essentiels des contributions qu'il contient. Il n'est pas possible de présenter ici les détails de la réalisation du projet ELICO, ni toutes les problématiques et les solutions liées à son thème central, la détermination. Nous nous contenterons donc d'indiquer les points les plus saillants.

1 Objectifs et problématique générale d'ELICO

Le but central du projet ELICO a été l'étude de l'évolution des déterminants du 13^{ème} au 18^{ème} siècle en français, en constituant et en exploitant un corpus annoté à cet effet. Du point de vue théorique, le projet rencontre trois grandes problématiques des sciences du langage : les systèmes de détermination, dont l'importance est centrale pour toutes les familles de langues, les facteurs qui influencent l'évolution des langues et le rapport entre théories linguistiques et observables linguistiques.

La problématique spécifique d'ELICO est de faciliter la synergie entre l'étude linguistique analytique, qui suppose une attention soutenue à la culture linguistique (théories, hypothèses, observations cruciales, représentations) et la prise en compte des observations de corpus, aussi bien les occurrences dans les textes que des propriétés plus globales, comme le genre, le domaine ou le style.

La configuration du projet a été déterminée par six facteurs :

- Choisir un domaine empirique déjà bien étudié aussi bien par les participants que dans la littérature linguistique, pour éviter de partir dans l'inconnu et pour pouvoir mettre en place rapidement des traits linguistiques fins,

¹ Le projet *Evolution linguistique et corpus* (ELICO) a bénéficié du soutien financier de l'Agence Nationale de Recherche française (ANR, projet 06-CORP-028-01).

- préférer des unités faiblement référentielles, dont la dépendance au domaine² soit moins forte que pour des noms ou des verbes par exemple,
- contribuer à l'étude d'une question linguistique actuelle et reconnue comme centrale dans une perspective descriptive, cognitive et typologique³, et ce quelles que soient les doctrines et les écoles, à savoir l'évolution du « système » de détermination,
- contribuer à l'articulation des communautés linguistiques dans leur diversité, notamment la linguistique théorique et la linguistique de corpus,
- exploiter l'expertise acquise pour des corpus électroniques d'Ancien Français par une équipe spécialisée, dont les solutions pratiques et méthodologiques peuvent être en partie recyclées (notamment voir BFM - Base de Français Médiéval, <http://bfm.ens-lsh.fr/>),
- faire du corpus, sous sa forme finale, non pas simplement une ressource, mais un véritable outil de travail, en prévoyant à la fois un balisage linguistique assez fin et une indication des différents traits textuels (« types » ou « genres » textuels, selon les terminologies).

Fondamentalement, ELICO utilise un balisage qui enregistre les occurrences de déterminants *et* les propriétés de leurs environnements locaux, par exemple le fait qu'ils apparaissent dans une question, dans une phrase qui est modalisée, dans un groupe nominal avec une tête nominale massique, etc. Pour ne pas imposer une analyse préalable des déterminants, l'annotation ne porte pas sur ceux-ci, mais sur un certain nombre de propriétés des environnements. Ce balisage permet de constituer un ensemble d'observables linguistiquement pertinents, puisqu'on classe les occurrences en fonction de traits linguistiques. Mais il permet également de mettre en relation les propriétés globales des textes (leur type textuel) avec des observables linguistiques déjà enrichis. Par exemple, on ne se demande pas simplement si tel type de texte contient significativement plus ou moins de *tout / toute / tous*, mais s'il contient significativement plus ou moins de *tout / toute / tous* dans certains environnements (les questions, les phrases négatives, etc.). L'étude de l'évolution des déterminants porte donc non seulement sur des aspects classiques dans les recherches sur l'évolution, c'est-à-dire la fréquence, l'apparition ou la disparition des unités, mais aussi sur leurs emplois et la sensibilité possible de ces emplois aux types de textes.

2 Détermination et évolution linguistique

Dans une langue comme le français, le mot *déterminant* renvoie en général aux articles simples (*le, un, cet*) mais aussi à des unités qui ne sont pas considérées comme des articles (*tout, certains, plusieurs, etc.*) et à des formes complexes comme *beaucoup de, un quelconque, un certain, tous les, n'importe quel, etc.* Les déterminants du français contemporain ont été considérés dans ELICO, de façon tout à fait traditionnelle, comme l'expression élémentaire qui, associée à un nom commun, permet d'obtenir un *Groupe Nominal* (GN en abrégé). Nous n'avons incorporé l'étude des *pronoms*, c'est-à-dire les unités qui peuvent remplacer un GN (*il, elle, celui-ci, le sien, etc.*), que ponctuellement pour les formes à emplois double.

L'importance des études consacrées aux déterminants dans les langues n'est plus à démontrer, aussi bien du point de vue de la quantité que de la variété des aspects pris en compte (voir par

² La dépendance au domaine correspond au fait que les domaines des différents textes peuvent influencer fortement la fréquence des noms, des verbes, des adjectifs, voire des prépositions et des constructions grammaticales (Resnik 1993).

³ On appelle *typologie* la partie de la linguistique qui s'occupe de la comparaison des langues et de l'évaluation de leurs analogies et de leur différences, cf. Greenberg (1974), Ramat (1987), Croft (2000), Son (2001) pour des introductions générales.

exemple Bach et al. (1995), Diessel (1999a,b), Haspelmath (1997), Lyons (1999), Himmelmann (2001), Landman (2004), Chesterman (2005), Dobrovie-Sorin & Beyssade (2005), Stark et al. (2007), Müller et Klinge (2008), Gomeshi et al. (2009)). Les déterminants sont des unités référentiellement assez pauvres. Elles ne désignent pas des objets, des événements ou des propriétés, et, de ce fait, se situent plutôt à l'extrémité de l'échelle qui ordonne lexical et grammatical. Toutefois, elles ont des conditions d'emploi complexes car elles reflètent plusieurs dimensions importantes de l'organisation grammaticale et discursive. D'abord, les déterminants touchent au problème général de la *quantification* (Peters & Westerståhl 2006), entendue comme l'ensemble des relations entre les différents domaines d'individus introduits par les termes référentiels (noms et verbes en particulier). Les contributions de Carlier et de Combettes contenues dans ce volume touchent à ce thème. Deuxièmement, les déterminants ont des propriétés anaphoriques qui contraignent leur insertion dans le discours et interfèrent avec l'organisation de celui-ci, voir la contribution de De Mulder et al. ici-même. Enfin, certains déterminants mobilisent des propriétés *modales* et *intersubjectives*, c'est-à-dire celles qui concernent la manière dont les agents de discours (locuteur, interlocuteur) voient les entités qui constituent la référence des noms déterminés. Par exemple, ces entités peuvent être vues comme connues ou non du locuteur, distinctes les unes des autres, équivalentes au regard de certaines propriétés, incorporées dans son domaine d'influence ou d'intérêt, etc. Les contributions de Bortolussi, Corblin, Jayez & Tovenà, et Pescarini concernent différents aspects de cette thématique.

D'un point de vue diachronique, les différentes facettes référentielles et textuelles des déterminants constituent un champ d'études difficile car s'y rencontrent une problématique très générale (en gros, l'anaphore) et la question de l'évolution des déterminants majeurs à partir du latin (voir Sabanéva 2003 pour leur origine latine). Rappelons, par exemple, que si les indéfinis communiquent qu'on introduit un référent situé relativement bas sur l'échelle des possibilités d'identification, les déterminants définis, démonstratifs ou possessifs communiquent au contraire que le référent peut être identifié, parce qu'il a été discursivement introduit, ou il fait partie de référents accessibles. Ici, une des questions principales est de savoir si les propriétés textuelles sont le reflet d'un profil sémantique propre à chaque déterminant, et qui trouverait à s'appliquer dans les textes, donc s'il y a une forte indépendance entre les propriétés sémantiques et textuelles, ou si, comme l'a proposé Guillot (2003, 2004) pour les démonstratifs, la fonction textuelle fait déjà partie du profil sémantique.

Des questions analogues se posent pour l'indéfini *uns* (Carlier 2001) et pour le défini (Epstein 1994, 1995, Carlier et De Mulder 2006). Rappelons que la classe empirique des indéfinis est quelque peu flottante; si les travaux linguistiques rangent les articles *un/des* et les numéraux dans cette catégorie, le statut d'unités comme *certain*s ou *tout* est bien moins clair. Cela tient à une tension entre un critère référentiel et un critère lié à la quantification généralisée; *certain*s et *tout* ne font pas référence à des individus particuliers et sont donc, sous cet angle, des indéfinis, voir Gondret (1976), Kleiber & Martin (1977), Corblin (2001), Jayez et Tovenà (2002, 2004), Schnedecker (2005). En revanche, *tout* n'est pas « symétrique »⁴, à la différence de *certain*s ou de *un* et, de ce point de vue, s'écarte de la majorité des indéfinis (Keenan 1987, Dobrovie-Sorin et Beyssade 2004, Tovenà et Jayez 1999a). Le groupe des indéfinis a des chevauchements avec deux autres groupes, à savoir les termes sensibles à la polarité et les termes de choix libre. D'un côté, les déterminants, et d'une manière générale, les éléments (déterminants, pronoms, GN figés) qui ont une affinité avec des environnements tels que la négation, les questions, etc. sont dits à *polarité négative* en linguistique (pour des synthèses

⁴ La symétrie correspond à la possibilité de permuter l'ordre des propriétés. La proposition *tout chat est un félin* (zoologiquement vraie) n'est pas équivalente à la proposition *tout félin est un chat* (zoologiquement fausse). Une remarque analogue vaut pour *chaque*. *Tout* et *chaque* continuent à être classés comme indéfinis dans certaines grammaires récentes (Riegel et al. 1994).

récentes, voir Giannakidou (1998), Tovenà (1998, 2001) et Tovenà et al. (2005) spécifiquement pour le français). L'origine et le comportement des déterminants à polarité négative, comme *le moindre, quelque ... que ce soit, un quelconque*, sont encore mal connus (voir Tovenà et Jayez (1999b) sur *le moindre*). De l'autre côté, certains déterminants sont liés à des états de croyance et/ou des états affectifs du locuteur, ce que l'on appelle des *attitudes* ou des *modalités* en sémantique. Par exemple, *un quelconque* correspond à un état d'ignorance du locuteur (*Marie a eu un problème quelconque* laisse penser que le locuteur ignore quel problème Marie a eu). L'étude de ces aspects est relativement récente et se relie à la fois à la question de la polarité et à celle des manifestations de la subjectivité dans la langue (voir Jayez et Tovenà (2002, 2005, 2006), Combettes (2004), Corblin (2004) et Corblin et al. (2010) pour quelques exemples récents sur le français).

Cette variété des perspectives possibles sur les déterminants se retrouve au niveau de l'expression de la *quantité*. Certains déterminants comme *beaucoup de, la plupart de, plusieurs* expriment ou paraissent véhiculer des indications de proportion. Si c'est assez clair pour *beaucoup* (voir Vogeleer (2003) pour sa sémantique fine), c'est moins évident pour *plusieurs*, dont l'analyse est liée au problème général des structures comparatives (Combettes 2006, Jayez 2006). D'une manière générale, il est important de rapprocher la notion de quantité à la fois des notions de mesure et d'argumentation, au sens de Ducrot.

Notre tâche a été de travailler sur la question générale de savoir comment a évolué, dans la période considérée, la gestion de la référence et la gestion de la cohésion textuelle par les déterminants. Cette question générale rencontre des questions plus spécifiques. Par exemple, comment évolue la distinction défini/indéfini, quel est le spectre des anaphores permises par les déterminants, quelle place occupent les démonstratifs ? Pour ce qui est des questions touchant à la polarité et aux attitudes, on ne disposait pas d'études diachroniques centrées sur ces problèmes, à la différence du point précédent. Le projet a contribué à combler ce manque et à rapprocher ainsi l'étude de l'évolution de la langue et des phénomènes reconnus comme très importants dans une perspective typologique. Les aspects abordés concernent notamment la nature des environnements de polarité (ont-ils évolué dans le temps?), le rapport entre polarité et sensibilité aux attitudes (les deux choses sont-elles liées, comme le soutiennent certaines approches actuelles?), et le rapport entre attitudes et non définitude (les déterminants sensibles aux attitudes sont-ils obligatoirement des indéfinis, comme le suggèrent, là encore, des approches actuelles?). Pour apprécier les changements linguistiques et en proposer des explications, nous prenons en compte deux aspects, les environnements linguistiques et les types textuels. C'est parce que cette deuxième facette constitue un critère possible de compétition⁵ que nous la retenons. Par exemple, pour exprimer une vérité générale sur une classe, un groupe, etc., on peut en français contemporain utiliser plusieurs déterminants (*le, les, un, tout*, etc.). Rien ne dit que ces déterminants se répartissent à peu près également selon les types de textes. Rien ne dit non plus que s'il y a égalité ou inégalité, elle est, pour les mêmes types de textes, constante dans le temps. Il est donc utile, lorsqu'on soupçonne une relation entre type de texte et emploi ou évolution d'un emploi, de pouvoir la mettre en évidence et comparer des explications possibles.

⁵ L'évolution des langues paraît dépendre en partie de phénomènes de *compétition*. A un moment donné, il existe, pour une classe de locuteurs, différents moyens d'exprimer le même contenu ou, plus exactement, le même type de contenu, avec des modulations.

3 Constitution du corpus et annotation

La collection de ELICO se compose de textes de différents genres couvrant six siècles (13^{ème} – 18^{ème}).⁶ Notre but initial était d'associer à chaque genre retenu un ensemble de textes de même volume global, mais nous avons dû assouplir cette contrainte face aux obstacles rencontrés, tels que, par exemple, les limites légales au droit d'exploitation des textes. En fonction de l'expérience acquise pour les projets BFM et « Déterminants modaux », nous avons limité ce volume à cent cinquante mille mots pour chaque demi-siècle, au départ, chiffre qui a été révisé à la baisse (quatre-vingt-dix mille mots par demi-siècle) pour les quatre premiers siècles. Nous sommes conscients que l'échelle adoptée pourrait pénaliser l'analyse des unités les moins fréquentes (pas assez d'attestations), mais une extension de la base est possible si le projet trouve une suite. Pour le présent, nous n'avons pas cherché à utiliser dans tous les cas des textes intégraux⁷ mais nous avons accordé la priorité au critère d'équilibre lorsque les conditions le permettaient. Une extension envisageable consiste à faire un premier relevé pour compter le nombre d'occurrences de chaque déterminant dans chaque genre et à enregistrer d'éventuelles sous/sur-représentations liées aux genres⁸. Cela nous permettrait de noter les dépendances entre genre et occurrences. L'exemple de *ledit*, associé aux textes juridiques (de Wolf 2003, Guillot et al. 2007, Mortelmans 2006) montre que c'est une donnée cruciale dans certains cas.

Nous avons travaillé sur 361 textes, en prose et en vers, représentatif des douze genres suivants : biographique, didactique, discours, économique, épistolaire, historique, juridique, littéraire, philosophique, politique, religieux et théâtral. A partir de cette sélection, nous avons constitué 435 extraits de trois mille mots chacun, en prenant soin de varier le point d'extraction pour les extraits qui ne sont pas des textes intégraux. Dans la mesure du possible, chaque extrait a été relu à partir d'une édition de référence pour corriger d'éventuelles erreurs. Tous ont été munis d'une description de base. Nous nous sommes appuyés principalement sur les descripteurs déjà utilisés pour la BFM (<http://bfm.ens-lsh.fr/>), dont nous avons adopté un sous-ensemble. Chaque extrait est associé à une fiche qui l'identifie et spécifie les informations suivantes :

auteur (nom)	
titre (titre)	
dates.auteur (deb,fin)	
#lieu.naissance (lieu)	champ facultatif
#lieu.composition (lieu)	champ facultatif
#dialecte (dialecte)	champ facultatif
date.manuscrit (date)	
date.edition (date)	
#editeur.commercial (éditeur)	champ facultatif

⁶ La quantité de textes antérieurs au 13^{ème} siècle est limitée et leur disponibilité pour une utilisation de domaine public est particulièrement réduite.

⁷ La notion même de « texte intégral » est loin d'être claire. Un chapitre d'ouvrage est un texte partiel par rapport à l'ensemble mais intégral par rapport à la division en chapitres. Comme cela est souvent signalé, un même texte peut comprendre des séquences textuelles de natures différentes, qui peuvent être perçues comme « intégrales » du point de vue du genre (cf. Adam 1992, 1999).

⁸ Techniquement, un genre induit une sous/sur représentation d'un phénomène par rapport à une mesure si la mesure indique que la répartition observée est significativement différente d'une répartition conforme à la distribution sur laquelle est fondée la mesure.

forme.texte (forme)	
#genre.texte (genre)	champ facultatif
source	

Pour ce qui est de la date, elle peut être soit une date unique, soit un intervalle, soit une séquence, notamment pour un ensemble de manuscrits. Voici par exemple, la fiche concernant un texte de François Villon.

```
<rdf:Description rdf:about="LeTestament_Villon">
  <rdf:type rdf:resource="&rens;Corpus_File"/>
  <rs:hasName>LeTestament_Villon</rs:hasName>
  <rs:hasAuthor rdf:resource="&aut;Villon_01-01-1431"/>
  <rs:hasTitle>Le testament</rs:hasTitle>
  <rs:hasEditionDate>1533</rs:hasEditionDate>
  <rs:CompositionDate>1456-1462</rs:hasCompositionDate>
  <rs:hasScientificEditor>Marot</rs:hasScientificEditor>
  <rs:hasTextForm>vers</rs:hasTextForm>
  <rs:hasTextGender>litteraire</rs:hasTextGender>
  <rs:has source>Internet Archive</rs:hassource>
</rdf:Description>
```

Il n'a pas été prévu d'étiquetage dans le cadre du projet, pour une raison pratique : concevoir un étiqueteur constitue un projet en soi, en utiliser un suppose des apprentissages différenciés pour les différentes tranches temporelles, ce qui est un peu lourd⁹ mais surtout risqué. Toutefois, dans la phase d'annotation, l'étiquetage de Tree Tagger (<http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>) est déclenché et son résultat proposé comme une aide à la décision, mais les informations constituées ne sont pas conservées dans l'annotation définitive.

Il est impossible de travailler sur la totalité (ensemble d'ailleurs difficile à définir) des déterminants. Le projet a choisi un premier ensemble de déterminants pour lequel l'interface d'annotation a été programmée. Les différentes formes qu'un déterminant peut avoir eues dans le temps, qui constituent donc ses « manifestations » en diachronie, ont été regroupées. Ces groupements, utilisés pour l'annotation, peuvent être utilisés aussi pour l'interrogation. Ensuite, un sous-ensemble de déterminants a été ciblé pour l'annotation. Le choix de déterminants retenus est le suivant (avec leurs variantes) : *le moindre, chaque, chaque un et tout, ledit, quelque, quelques et quelque un, aucun et aucuns, plusieurs, maint et moult*.

Le résultat de l'annotation est une collection d'environ 20000 fiches qui, associée à l'information sur les textes, constituent la base de ELICO. Ces fiches comportent les informations suivantes.

⁹ Lourd parce que, quelles que soient les qualités intrinsèques des étiqueteurs connus tels que CORDIAL ou TREETAGGER, le fait qu'ils ne soient pas interactifs (ne permettent pas de modifier leur vocabulaire avant de déclencher un étiquetage) les pénalise inévitablement dans un certain nombre de situations.

- Il y a d'abord un bloc d'informations sur l'occurrence elle-même : 1. la forme de cette occurrence, 2. son type syntaxique (déterminant/pronom/ adverbe associé à un déterminant), 3. la présence de *de* dans la forme, et 4. l'insertion dans un multidéterminant.

- Ensuite, il y a des traits relatifs au Nom (ou GN) attaché au déterminant, en premier lieu 5. la forme du nom. Suivent des informations grammaticales telles que 6. le nombre 7. le genre et 8. la présence d'une coordination nominale dans le GN (det N1 coord N2). Puis on cerne les propriétés à fort impact sémantique, telles que 9. le caractère abstrait du nom (événement, action, sentiment, qualité), le caractère concret, le caractère « unité de temps » ou instantiation/mesure, 10. le caractère massif ou comptable, 11. le statut de nom propre ou de nom commun, 12. la présence de modifieur(s) (adjectif, participe, subordonnée relative, complément du nom), 13. la forme du/des modifieur(s) et 14. sa/leur place(s) par rapport au nom.

- Enfin, il y a les traits relatifs au contexte : le trait 15., statut grammatical du GN, décrit la fonction du GN qui contient le déterminant dans la phrase. Il y a trois traits relatifs à la forme verbale de la clause qui contient le groupe nominal, à savoir 16. mode du verbe de la clause 17. temps, 18. voix. Enfin, on annote de l'information relative à la clause : 19. verbe principal modal, 20. nature de la phrase (affirmative, négative, interrogative, interro-négative, exclamative) et, principalement pour calculer des effets de double négation, 21. le fait que *sans* prend le GN dans sa portée, et 22. la présence d'une structure conditionnelle.

4 Interrogation

L'interrogation se fait à travers une interface spécifique, distincte de l'interface d'annotation. Elle permet de récupérer toutes les attestations obéissant aux critères choisis par l'utilisateur parmi un ensemble de critères proposés par l'interface. D'un côté, il est possible de récupérer toutes les formes diachroniques d'un déterminant avec une requête simple portant sur le groupement en entier. D'un autre côté, on peut définir des critères spécifiques de recherche. A titre d'exemple, nous montrons ici une partie de la sélection de critères. Dans cet exemple, l'utilisateur a demandé les formes de *quelque* qui figurent dans une phrase à l'indicatif avec une marque négative. Les valeurs des autres traits contextuels sont laissées indéterminées.

The screenshot shows a search interface titled "Traits relatifs au contexte" with a dropdown arrow. It lists several criteria with corresponding input fields:

- 15. Statut grammatical du groupe nominal : [empty dropdown]
- 16. Mode : [indicatif dropdown]
- 17. Temps : [empty dropdown]
- 18. Voix : [empty dropdown]
- 19. Verbe principal modal : [empty dropdown]
- 20. Nature de la phrase : [negative dropdown]
- 21. Dans la portée de «SANS» : [empty dropdown]
- 22. Présence d'une structure conditionnelle : [empty dropdown]

At the bottom left, there is a "Soumettre" button.

L'exploration fournit 7 réponses, dont un texte de Marot, *Le voyage de Gênes*. Les propriétés de l'occurrence sont décrites ci-dessous

quelque --- #LeVoyageDeGenes_Marot V ^	
Type : déterminant	
Forme avec "DE" : non	
Multidéterminant : non	
<hr/>	
Traits relatifs au nom (ou GN) attaché au déterminant	
Nom : macule	
Genre : féminin	
Nombre : singulier	
Coordination nominale du GN : non	
<hr/>	
Marquage sémantique de l'occurrence du N	
Nom abstrait nom concret unité de temps instanciation autres : autre	
Nom massif nom comptable : comptable	
Nom propre nom commun : commun	
Présence d'un modifieur auprès du nom : non	
Modifieur(s) :	
Place du modifieur par rapport au nom : non	
<hr/>	
Traits relatifs au contexte	
Statut grammatical du groupe nominal : CD	
Mode : indicatif	
Temps : passe	
Voix : active	
Verbe modal principal : non	
Nature de la phrase : negative	
Dans la portée de "SANS" : non	
Présence d'une structure conditionnelle : non	

On peut avoir accès à un contexte limité (pour des raisons légales), qui, dans cet exemple, est le suivant (l'occurrence décrite est en gras souligné).

et banieres Povres souldars sortent de leurs tesnieres, Courent les rues comme demoniacles, Gastans des nobles tours, maisons et pinacles. Lors eussiez veu, contre murs et pallis, Oû imprimées estoyent les fleurs de lis, Lascher leurs traitz ; les ungs les derompoient Les aultres fange à l'encontre jettoient, En tel façon qui n'en demoura nulle Qui fust entiere ou n'eust **quelque** macule. Durant ce trouble, ceste turbe maligne Va conspirer d'aller mettre en ruine Le Castellat et de bouter à mort Tous les François, qui lors gardoyent le fort, Lesquelz, sachans la place estre non forte, Et d'aultre part voyans ceste cohorte Plus animez à faire leur emprinse Que juifz à faire de Jhesucrist la prinse, Delibererent de leur rendre la place En leur

Cet extrait illustre le problème de l'interaction avec la négation dans le cadre d'une annotation contextuelle. Certains éléments, dont *quelque*, sont réputés peu compatibles avec une négation dans la même clause (?? *Il n'a pas fait quelque effort* vs *Il n'a fait aucun effort*). Toutefois, la présence de deux négations améliore en général sensiblement la perception des exemples (*non sans quelque effort*). C'est le cas ici, où on trouve une occurrence de *quelque* dans une clause négative *mais* elle-même enchâssée sous une autre négation. Un usage raisonné de l'annotation suppose que l'on regarde le contexte élargi avant de tirer des conclusions hâtives, le contexte immédiat ne servant que de signal.

5 Les articles de ce volume

Dans le projet ELICO, plutôt que de viser une cartographie ponctuelle des déterminants, les diverses recherches ont été centrées autour des trois questions suivantes : les propriétés référentielles et textuelles des déterminants, les quantificateurs, et l'interaction avec la polarité et avec les attitudes du locuteur. Les articles dans ce volume sont représentatifs de ces trois domaines.

5.1 Les démonstratifs

L'objectif de l'article « *Ce N-ci* et *ce N-là* en moyen français » de Walter De Mulder, Céline Guillot et Jesse Mortelmans est de décrire l'évolution de la valeur sémantique des démonstratifs « complexes » *ce N-ci* et *ce N-là* à partir de leur apparition. L'attention des auteurs porte surtout sur la valeur des éléments *-ci* et *-là*, partant de l'hypothèse que la valeur

du démonstratif simple (*ce*) telle qu'elle a été décrite dans la littérature se maintient dans les formes complexes. Après avoir analysé les occurrences de *ce N-ci* et *ce N-là* dans un corpus de textes en moyen français, les auteurs s'attaquent à la question de savoir si ces formes complexes ont repris la répartition des valeurs sémantiques des paradigmes *cist* et *cil* en ancien français.

5.2 Quantification

Beaucoup a été analysé comme adverbe, comme déterminant-adverbe ou comme expression catégoriellement sous-spécifiée. Dans « De *multum* à *beaucoup* : entre adverbe et déterminant nominal », Anne Carlier aborde la question de la catégorisation morphosyntaxique de *beaucoup* dans une perspective diachronique en mettant en évidence deux parallélismes entre *multum* en latin classique et *beaucoup* tel qu'il se grammaticalise en moyen français : (i) ils peuvent avoir dans leur portée un nom ou un verbe, mais non un adjectif ou un adverbe ; (ii) ils se rapportent à l'objet direct (exprimé en latin par l'accusatif non régi par une préposition) ou au sujet (au cas nominatif) de certains verbes et ne sont pas attestés dans d'autres positions syntaxiques. Ces parallélismes sont expliqués à partir de l'origine nominale de *multum* et *beaucoup*. En tant que formes nominales, *multum* et *beaucoup* sont utilisés pour quantifier l'argument interne du verbe, donc comme capables de quantifier le procès verbal en tant que tel. Ils peuvent ensuite devenir des adverbes à part entière sans être associés à une position argumentale. Ce n'est qu'à un stade ultérieur qu'ils peuvent quantifier un syntagme nominal qui ne soit plus intégré dans le prédicat verbal et qu'ils deviennent ainsi déterminants nominaux à part entière. Leur statut hybride, entre adverbe et déterminant nominal, se maintient quand ils sont associés à l'argument interne du verbe. L'éclairage diachronique qu'apporte cette étude permet de mieux comprendre le fonctionnement de *beaucoup* en français moderne : (i) sur le plan distributionnel, il permet de comprendre pourquoi *beaucoup* en tant qu'adverbe se rapporte à un verbe, mais non à un adjectif ou un adverbe ; (ii) il apporte des précisions sur la catégorisation morphosyntaxique de *beaucoup* dans ses différents emplois en français moderne et éclaire des différences syntaxiques par rapport à des déterminants quantifiants comme *plusieurs* ; (iii) du point de vue sémantique enfin, il permet de comprendre dans quels contextes *beaucoup* se comporte comme un quantifieur ordinaire, pouvant entrer dans des relations de portée.

Le processus de grammaticalisation de *plusieurs* est au centre de l'article « Aspects de la grammaticalisation d'un déterminant: le cas de *plusieurs* » de Bernard Combettes. Ce processus de grammaticalisation, au plan morphosyntaxique, a conduit la forme *plusieurs* du statut adjectival au statut de déterminant nominal et, au plan sémantique, a, d'une part, limité la valeur de cet indéfini à l'expression de la quantité faible et l'a, d'autre part, doté d'une orientation à partir de l'unité. Après avoir rappelé l'origine comparative de *plusieurs*, origine qui permet d'expliquer le type d'évolution qui a affecté cette expression, l'auteur examine les différents contextes qui ont pu jouer un rôle dans les changements successifs, en observant surtout les textes du moyen français. C'est en effet durant cette période que semble disparaître peu à peu l'emploi adjectival et se préciser la valeur sémantique moderne. Cette évolution de *plusieurs*, que l'on peut considérer comme une spécialisation, syntaxique et sémantique, et qui fait que cet indéfini s'oppose par exemple de plus en plus nettement à *quelques*, apparaît comme une réalisation particulière d'une tendance de fond qui affecte d'autres catégories grammaticales du français

5.3 Indéfinis et formes de déterminants modaux

Certains des travaux groupés dans cette section tournent autour du déterminant *quelque*. Dans sa forme au pluriel, il est souvent comparé à *plusieurs* pour indiquer une quantité modeste. Dans sa forme au singulier, il est aujourd'hui ressenti comme littéraire dans la plupart de ses

emplois, mais il a des particularités qui dépassent largement celles d'un simple indéfini et qui constituent un défi pour les linguistes. Les termes latins recouvrant les emplois du français *quelque(s)* sont discutés par Bernard Bortolussi dans son article « Quelques *quelque(s)* en latin ». Ces termes latins sont très variés et dépendent étroitement des contextes syntaxiques et pragmatiques. Trois formes de pluriel : *aliquot*, *quidam* et *aliqui* se répartissent les emplois de *quelques*. De même, trois formes de singulier *quidam*, *aliquis* et *quis* renvoient respectivement au [spécifique, connu], au [spécifique, inconnu] et au [non spécifique]. *Quis* s'oppose également à un autre indéfini, *quisquam*, en contexte négatif. L'évolution en latin tardif montre une tendance à la confusion entre les différents indéfinis précédents, ainsi qu'une concurrence avec le distributif *quisque* et les indéfinis de choix libre, dont *qualiscumque/qualisque*.

Dans « Description et évolution de *quelque* », Jacques Jayez et Lucia M. Tovena proposent une caractérisation sémantique de *quelque* qu'ils soumettent à l'épreuve des données sur l'évolution pour tester sa compatibilité avec les faits et sa capacité explicative. Leur hypothèse de travail est que *quelque* est un indéfini qui repose sur l'inférence et l'ignorance. La forme *quelque NP*, où *P* est une propriété, signale que l'existence d'une entité de type *N* satisfaisant *P* est une information obtenue par inférence (aspect évidentiel) et que l'identité exacte de cette entité demeure inconnue (aspect épistémique). Les auteurs montrent que le rapprochement intuitivement évident entre *quelque* et les déterminants dits à choix libre, par exemple *n'importe quel* ou *un quelconque*, est plausible à condition de faire intervenir l'aspect évidentiel. Une attention particulière est portée à la difficulté de *quelque* à se trouver dans la portée immédiate de la négation et elle est analysée comme le reflet de l'organisation sémantique propre de ce déterminant, c'est-à-dire du fait qu'il véhicule à la fois un contenu principal, sa valeur existentielle d'indéfini, et une implicature conventionnelle, les valeurs inférentielles et épistémiques. Enfin, quelques perspectives d'analyse sur l'origine et l'évolution de *quelque* sont mises en place et ceci permet aux auteurs d'identifier un certain nombre de problèmes importants, parmi lesquels la question de savoir si c'est la concessivité qui dérive de l'ignorance ou l'inverse et la question de la possibilité ou de la manière de réconcilier le sens concessif avec l'interprétation existentielle.

On retrouve la question de la difficulté de *quelque* à se trouver dans la portée immédiate de la négation au cœur de l'article « La préférence existentielle du déterminant *quelque* » de Francis Corblin. Cet article montre que les pronoms indéfinis (*quelqu'un*, *quelque chose*) du français ne peuvent pas être analysés comme des termes sensibles à la polarité positive, ni comme doublement sensibles à la polarité négative. Selon cet auteur, les restrictions de portée concernant la négation et le non-fonctionnement dans les énoncés génériques de ces items ne peuvent se justifier qu'en leur prêtant une vocation pour l'existentiel. Cette position serait un retour partiel à l'approche initiale de Russell (1905), qui avait proposé une analyse sémantique de l'anglais *some* qui revient à le traiter comme un quantificateur existentiel. L'article établit que cette vocation est une propriété lexicale du déterminant *quelqu-* du français.

Enfin, Sandrine Pescarini, dans son article « *N'importe quel* en perspective diachronique » se propose de présenter la construction et le développement des indéfinis formés sur le verbe *importer*, notamment *n'importe quel*. Les éléments de cette famille sont employés régulièrement depuis le début du XIX^e siècle. Ils ont d'abord été construits de manière compositionnelle. Suite au processus de lexicalisation, on ne considère plus actuellement le sens de tous les constituants. Les trois valeurs interprétatives attribuées à *n'importe quel* : sont l'élargissement, la dépréciation et l'indifférence, valeurs typiques des termes de choix libre, dont l'auteur montre la filiation par rapport à l'origine diachronique. Outre son intérêt intrinsèque, cette contribution permet de voir comment une forme, dont la motivation est propre au français, a des valeurs synchroniques très proches de l'anglais *any*, qui a été

considéré dans la littérature sémantique comme une sorte de prototype des termes de choix libre.

Références

- Adam, J-M. (1992). *Les textes, types et prototypes : récit, description, argumentation, explication et dialogue*. Paris : Fernand Nathan.
- Adam, J-M. (1999). *Linguistique textuelle. Des genres de discours aux textes*. Paris : Nathan.
- Bach, E., Jelinek, E., Kratzer A. et Partee, B. H. (éds) (1995). *Quantification in Natural Languages*. Dordrecht : Kluwer.
- Blutner, R. et Zeevat, H. (2004). *Optimality Theory and Pragmatics*. Basingstoke et New York : Palgrave Macmillan.
- Brinton, L. J. et Traugott, E. C. (2005). *Lexicalization and Language Change*. Cambridge : Cambridge University Press.
- Bybee, J. L. et Hopper, P. J. (2001). *Frequency and the emergence of linguistic structure*. Amsterdam : Benjamins.
- Carlier, A. (2001). La genèse de l'article *un*. *Langue Française* 130, 65-88.
- Carlier A. (2004). Sur les premiers stades de développement de l'article partitif, *Scolia* 18, 117-147.
- Carlier, A. et De Mulder, W. (2006). Les premiers stades de développement de l'article défini. *Verbum*, 81-110.
- Carlier, A. et Goyens, M. (1998). De l'ancien français au français moderne: régression du degré zéro de détermination et restructuration du système des articles. *Cahiers de l'Institut de Linguistique de Louvain* (Louvain-la-Neuve; internationaal tijdschrift) 24, 77-112.
- Chesterman, A. (2005). *On Definiteness*. Cambridge : Cambridge University Press.
- Combettes, B. (2001). L'émergence d'une catégorie morphosyntaxique : les déterminants du nom en français. *Linx* 45, 117-126.
- Combettes, B. (2004). *Quelque* : aspects diachroniques. *Scolia* 18, 9-40.
- Combettes, B. (2006). *Plusieurs* : étude diachronique. *Scolia* 20.
- Corblin, F. (1987). *Indéfini, défini et démonstratif. Constructions linguistiques de la référence*. Genève : Droz.
- Corblin, F. (1997). Les indéfinis : variables et quantificateurs. *Langue Française* 116, 8-32.
- Corblin, F. (2001). Où situer certains dans une typologie des groupes nominaux? Dans : Georges Kleiber, Brenda Laca et Liliane Tasmowski (dir.), *Typologie des groupes nominaux*, Rennes : Presses Universitaires de Rennes, 99-117.
- Corblin, F. (2004). *Quelque*. Dans Francis Corblin et Henriëtte de Swart (éds), *Handbook of French Semantics*, Stanford : CSLI Publications, 99-109.
- Corblin, F., Tovina, L. M. et Vlachou E. (2010). *Les indéfinis de choix libre du français*, *Langue Française* Numéro thématique 166.
- Croft, W. (2000). *Explaining linguistic Change – an Evolutionary Approach*. Londres : Longman.
- Cornish, F. (1999). *Anaphora, Discourse and Understanding. Evidence from English and French*. Oxford : Clarendon Press.
- De Mulder, W. (1992). Il y a sens et signification. La sémantique et pragmatique des démonstratifs et définis. Thèse de doctorat, Université d'Anvers.
- De Mulder, W. (1997). Les démonstratifs : des indices de changement de contexte. Dans Nelly Flaux, Danièle Van de Velde et Walter de Mulder (éds), *Entre général et particulier : les déterminants*, Arras : Presses de l'Université d'Artois, 137-200.
- Diessel, H. (1999a). *Demonstratives. Form, Function, and Grammaticalization*. Amsterdam : John Benjamins.

- Diessel, H. (1999b). The morphosyntax of demonstratives in synchrony and diachrony. *Linguistic Typology* 3, 1-49.
- Dobrovie-Sorin, C. et Beyssade, C. (2004). *Définir les indéfinis*. Paris: CNRS Éditions.
- Epstein, R. (1994). The Development of the Definite Article in French. Dans: W. Pagliuca (éd.), *Perspectives on Grammaticalization*. Amsterdam-Philadelphie : John Benjamins , 63-80.
- Epstein, R. (1995). The later stages in the development of the definite article : evidence from French. Dans H. Andersen (éd.), *Historical Linguistics 1993*, Amsterdam-Philadelphie: John Benjamins, 159-176.
- Giacalone-Ramat, A. & Hopper, P. (éds) (1998). *The Limits of Grammaticalization*. Amsterdam-Philadelphie : John Benjamins.
- Giannakidou, A. (1998). *Polarity Sensitivity as (Non) Veridical Dependency*. Amsterdam: John Benjamins.
- Gomeshi, Jila, Paul, Ileana et Wiltschko M. (éds) (2009). *Determiners*. Amsterdam: John Benjamins.
- Gondret, P. (1976). *Quelques, plusieurs, certains, divers: étude sémantique*. *Le Français Moderne* 2, 143-152.
- Greenberg, J.H. (1974). *Language Typology : A Historical and Analytic Overview*. La Haye : Mouton.
- Guillot, C. (2003). Le rôle du démonstratif dans la cohésion textuelle au XV^{ème} siècle. *Éléments de grammaire textuelle*. Thèse de doctorat, ENS-LSH, Lyon.
- Guillot, C. (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.
- Guillot, C., Heiden, S. et Lavrentev, A. (2007). Typologie des textes et des phénomènes linguistiques pour l'analyse du changement linguistique avec la Base de Français Médiéval. Dans Malrieu, D. (dir.), *Corpora et Questionnements du littéraire, Actes des Journées d'étude internationales, LINX*, numéro spécial, p. 125-139.
- Haspelmath, M. (1997). *Indefinite Pronouns*. Oxford : Clarendon Press.
- Himmelmann, N. (2001). Articles. Dans Haspelmath, Martin, König, Ekkehard, Oesterreicher, Wulf et Raible, Wolfgang (éds), *Language Typology and Language Universals*, vol. 1, Berlin et New York : de Gruyter,
- Hopper, P. J. & Traugott, E. C. (2003²/1993). *Grammaticalization*. Cambridge University Press.
- Jayez, J. (2006). How many are « several »? *Belgian Journal of Linguistics* 19.
- Jayez, J. et Tovenà, L. M. (2002). Determiners and (Un)certainity, *Proceedings of SALT 12*, CLC Publications, 164-183.
- Jayez, J. et Tovenà, L. M. (2004). *Tout* as a genuine free-choice item. Dans Francis Corblin et Henriëtte de Swart (éds), *Handbook of French Semantics*, Stanford : CSLI Publications, 71-81.
- Jayez, J. et Tovenà, L. M. (2006). Epistemic determiners, *Journal of Semantics* 23, 217-250.
- Keenan, E. L. (1987). A semantic definition of « Indefinite NP ». Dans Eric J. Reuland et Alice G.B. ter Meulen, *The Representation of Indefiniteness*, Cambridge (MA) : MIT Press, 286-317.
- Kleiber, G. et Martin, R. (1977). La quantification universelle en français. *Semantikos* 2, 19-36.
- Landman, F. (2004). *Indefinites and the Type of Sets*. Oxford: Blackwell Publishing.
- Lyons, C. (1999). *Definiteness*. Cambridge : Cambridge University Press.
- Mortelmans, J. (2006). *Ledit vs le démonstratif en moyen français : quels contextes d'emploi?* *Langue Française* 152, 70-81.

- Müller, H. H. et Klinge A. (éds) (2008). *Essays on nominal determination*. Amsterdam: John Benjamins.
- Peters, S. et Westerståhl, D. (2006). *Quantifiers in Language and Logic*. Oxford : Oxford University Press.
- Ramat, P. (1987). *Linguistic Typology*. New York, Berlin & Amsterdam: Mouton de Gruyter.
- Resnik, P. (1993). Selection and Information: A Class-Based Approach to Lexical Relationships. Thèse de doctorat, Université de Pennsylvanie.
- Riegel, M., Pellat, J-C. et Rioul, R. (1994). *Grammaire méthodique du français*. Paris : Presses Universitaires de France.
- Russel, B. (1905). On denoting, *Mind* 14, 479-493.
- Schnedecker, C. (2005). *Certain* et ses avatars (*certain N / un certain N ; certains N / de certains N ; certains*) : approche diachronique. *Travaux de Linguistique* 50, 131-150.
- Son, J. J. (2001). *Linguistic Typology. Morphology and Syntax*. Harlow : Pearson Education.
- Stark, E., Leiss E. et Abraham W. (éds) (2007). *Nominal determination*. Amsterdam: John Benjamins.
- Tovena, L. M. (1998). *The Fine Structure of Polarity Sensitivity*. New York : Garland.
- Tovena, L. M. (2001). The phenomenon of polarity sensitivity: questions and answers, *Lingua e stile* XXXVI:1, 131-167.
- Tovena, L. M., Déprez, V. et Jayez, J. (2005). Polarity sensitive items, Dans F. Corblin and H. de Swart (éds.) *Handbook of French semantics*, CSLI Publications, 391-415.
- Tovena, L. M. et Jayez, J. (1999a). Déterminants et irréférence. L'exemple de *tout*. Dans *Référence nominale et temporelle*, Jacques Moeschler et Marie-José Reichler-Béguelin (éds) Berne : Peter Lang, 35-268.
- Tovena, L. M. et Jayez, J. (1999b). *Any* : from scalarity to arbitrariness. Dans Corblin, Francis, Dobrovie-Sorin, Carmen et Marandin, Jean-Marie (éds), *Empirical Issues in Formal Syntax and Semantics 2*, La Haye : Thesus, 39-57.
- Traugott, E. C. et Dasher, R. B. (2002). *Regularity in Semantic Change*. Cambridge: Cambridge University Press.
- Traugott, E. C. et Heine, B. (1991). *Approaches to grammaticalization* (2 vol.). Amsterdam-Philadelphia : John Benjamins Publishing Company.
- Vogeleer, S. (2003). Les quatre lectures du quantificateur *beaucoup de*. *Langages* 151, 43-65.
- de Wolf, A. (2003). Un nouveau déterminant : Le déterminant anaphorique *-dit* en français médiéval. *Verbum* 25, 335-351.