

CORPUS **Corpus**
9 | 2010
La syntaxe de corpus

Quand un corpus rencontre un adjectif du troisième type. Etude distributionnelle de *prochain*

Christophe Benzitoun, Solène Bresson, Laure Budzinski, Jeanne-Marie Debaisieux et Klara Holzheimer



Édition électronique

URL : <http://journals.openedition.org/corpus/1927>

ISSN : 1765-3126

Éditeur

Bases ; corpus et langage - UMR 6039

Édition imprimée

Date de publication : 1 novembre 2010

Pagination : 245-264

ISSN : 1638-9808

Référence électronique

Christophe Benzitoun, Solène Bresson, Laure Budzinski, Jeanne-Marie Debaisieux et Klara Holzheimer, « Quand un corpus rencontre un adjectif du troisième type. Etude distributionnelle de *prochain* », *Corpus* [En ligne], 9 | 2010, mis en ligne le 06 juillet 2011, consulté le 23 avril 2019. URL : <http://journals.openedition.org/corpus/1927>

Quand un corpus rencontre un adjectif du troisième type. Etude distributionnelle de *prochain*

Christophe BENZITOUN, Solène BRESSON,
Laure BUDZINSKI, Jeanne-Marie DEBAISIEUX,
Klara HOLZHEIMER
Nancy Université & ATILF CNRS

1. Introduction

Nous nous sommes intéressés à l'adjectif *prochain* pour diverses raisons. En premier lieu, pour son comportement singulier. Cette singularité tient au fait qu'il n'a ni les propriétés d'un adjectif qualificatif, ni celles d'un adjectif relationnel. Il peut très difficilement apparaître en position d'attribut (**Cette semaine est prochaine*) ou être modifié par *très*, contrairement aux qualificatifs, et il n'est pas morphologiquement dérivé d'un substantif, contrairement aux relationnels. C'est la raison pour laquelle Schnedecker (2002a) le classe parmi les adjectifs « du troisième type ».

Par ailleurs, l'adjectif *prochain* n'a fait l'objet d'aucune étude systématique. Seuls quelques auteurs s'y sont intéressés (Schnedecker 2002a ; Berthonneau 2002), mais dans le cadre d'études essentiellement basées sur l'introspection et la prise en compte de quelques rares énoncés attestés. Nous ne disposons ainsi d'aucunes indications sur la fréquence de *prochain* ni sur celle de ses différents emplois. Dans la présente étude, nous nous proposons d'enrichir la description de *prochain* en nous intéressant plus particulièrement à son utilisation en français parlé.

Pour ce faire, nous prendrons en compte les paramètres qui nous ont paru avoir une incidence sur le fonctionnement de cet adjectif. Certains paramètres ont déjà été mentionnés par les auteurs pré-cités, d'autres ont été mis en évidence au travers de

la description des usages attestés en corpus. D'un point de vue distributionnel, le nombre (singulier/pluriel), la place de *prochain* par rapport à son nom recteur ainsi que le nom recteur lui-même sont de première importance. De même, la dimension du « genre de texte » doit également être prise en compte. Ainsi, la partie orale sur laquelle a porté notre étude présente une répartition plus nette que l'écrit, ce qui montre la pertinence de travailler sur de telles données.

Cette démarche permet une comparaison fructueuse entre nos résultats et la description proposée par Berthonneau (2002) et illustre la problématique de l'articulation entre linguistique sur corpus et linguistique introspective. Nous proposerons, en outre, une réflexion sur les descriptions proposées dans les entrées de dictionnaires consacrées à cet adjectif. Nous verrons que le travail sur corpus permet d'enrichir et compléter les sources actuelles.

La première partie présente le corpus sur lequel est basée notre étude ainsi que la démarche que nous avons suivie. Nous détaillerons ensuite les résultats les plus significatifs de l'étude avant de nous intéresser aux caractéristiques des noms recteurs de *prochain*. Nous terminerons par une brève comparaison entre les usages attestés en corpus et les entrées de dictionnaire.

2. Présentation des corpus et des outils utilisés

Le corpus utilisé est composé de transcriptions d'oral et de textes écrits. Par commodité, nous emploierons, dans la suite de l'article, le terme « oral » pour désigner la portion d'oral sur laquelle nous travaillons et nous en ferons de même pour l'écrit. Cette commodité de dénomination ne doit pas être interprétée comme une opposition simpliste entre oral et écrit. Nos conclusions ne sont valables que pour les données observées et non pas pour le français parlé en général. Nous avons privilégié les conversations spontanées dans les corpus oraux et avons pris des textes diversifiés de style élaboré pour l'écrit.

Oral (environ 2 millions de mots) : *Corpus de Français Parlé Parisien* (CFPP2000), *Corpaix* (version de mai 2000),

*Quand un corpus rencontre un adjectif du troisième type.
Etude distributionnelle de prochain*

Corpus de Référence du Français Parlé, ainsi que la partie de la base *Phonologie du Français Contemporain* contenant des conversations libres enregistrées en France.

Écrit (environ 2 millions de mots) : Presse (*Courrier international, Le Monde, Le Monde diplomatique, le Nouvel Observateur*) ; discours politiques (J. Chirac, L. Jospin, F. Mitterrand) ; textes scientifiques (ouvrages publiés aux éditions du CNRS, revues parues chez Hermès, *Pour la science, Sciences et avenir*) ; textes institutionnels (Assemblée Nationale, textes juridiques et législatifs) ; autres textes divers (philosophie, critiques littéraires, critiques cinéma, nouvelles de science fiction, romans).

Le corpus écrit provient du *Corpus Evolutif de Référence du Français* (CERF), constitué à l'Université de Provence sous la direction de Jean Véronis.

Les concordances ont été extraites à l'aide du concordancier *Contextes* développé à l'Université de Provence par Jean Véronis. Elles se présentent sous la forme d'un tableau ayant pour première colonne le nom de la division d'où est extraite l'occurrence suivie du contexte gauche, du mot cible, du contexte droit ainsi qu'une clef de tri gauche. Après obtention des concordances, l'étude distributionnelle a été réalisée sous Excel grâce notamment aux fonctions de *tris* et de *filtres*. Cinq colonnes supplémentaires ont été ajoutées contenant : la position de *prochain* par rapport à son nom recteur, le genre¹ (masculin / féminin), le nombre (singulier/pluriel), la forme du déterminant ainsi que le nom recteur et son sens simplifié, en particulier lorsque ce dernier est un temporel. Chaque ligne de concordance a fait l'objet d'un étiquetage manuel en fonction des cinq paramètres décrits ci-dessus, lequel étiquetage a été exploité grâce à la fonction de filtre. Les exemples du type *aimer son prochain* et *à la prochaine* ont été écartés. Il s'agit de fonctionnements particuliers ne comportant pas de noms recteurs, et pour lesquels on relève 11 occurrences à l'écrit et 5 à l'oral. Oral et écrit ont été analysés dans deux fichiers séparés.

¹ Ce paramètre s'est finalement avéré peu pertinent, raison pour laquelle nous n'en parlons pas par la suite.

Lorsque cela s'avérait intéressant, nous avons pris en compte des données supplémentaires. Dans cette perspective, nous avons choisi de calculer automatiquement les fréquences en extrayant le premier mot directement à droite et à gauche de *prochain*, étant donné la quantité d'occurrences à traiter (1662). Puis nous avons lemmatisé (à l'aide du logiciel *Cordial Analyseur* de la société *Synapse Développement*) ces deux listes de mots pour tester l'influence de ce paramètre sur nos résultats. Nous avons ensuite utilisé un logiciel qui dénombre les occurrences (*Dico*, développé par Jean Véronis) et supprimé manuellement les cas qui n'étaient pas des noms. Pour chaque position (antéposé et postposé), nous disposons donc de deux listes : la liste des formes et la liste des lemmes des noms recteurs. Nous sommes conscients que par ce procédé nous n'avons pas récupéré l'ensemble des noms recteurs. Toutefois, nous avons observé suffisamment d'exemples pour savoir que le nom recteur se trouve à une écrasante majorité à une distance $n+1$ ou $n-1$ de *prochain*.

Notre approche combinant outils informatiques et interventions manuelles peut surprendre. La part laissée à l'intervention manuelle peut d'ailleurs paraître démesurée au regard des fonctionnalités des logiciels actuels. Mais en parallèle du présent travail, nous menons une réflexion sur l'automatisation des tâches et souhaitons évaluer et comparer des résultats obtenus en fonction de différents degrés d'automatisation. Nous avons utilisé ici la méthode qui nous semblait fournir le meilleur compromis entre le temps de travail et le degré de confiance que l'on pouvait accorder aux résultats.

3. Principaux résultats de l'analyse distributionnelle

L'étude distributionnelle met au jour une forte sélection lexicale entre le nom recteur et *prochain*. En effet, *prochain* est très majoritairement accompagné de noms temporels. Ces derniers représentent d'ailleurs les deux tiers des emplois observés et même 80 % des emplois à l'oral ! Cette sélection lexicale est fortement conditionnée par la position de l'adjectif par rapport à son nom recteur, le nombre ainsi que le type de corpus. Ces contraintes d'emploi ont été peu détaillées dans les travaux

*Quand un corpus rencontre un adjectif du troisième type.
Etude distributionnelle de prochain*

antérieurs. Neutraliser ne serait-ce que l'un de ces paramètres donnerait une vision très parcellaire voire erronée du fonctionnement de cet adjectif. Par exemple, le fait de travailler à partir des lemmes et non des formes « brutes » présentes dans le corpus initial masque la corrélation entre le nombre et la position relative. De même, l'oral propose une distribution d'une régularité saisissante qui ne peut être saisie qu'au travers de la distinction singulier / pluriel. Nous rejoignons donc la prudence manifestée par J. Sinclair à l'égard des corpus « enrichis » :

In corpus-driven linguistics you do not use pre-tagged text, but you process the raw text directly and then the patterns of this uncontaminated text are able to be observed. (Sinclair, 2004 : 191)

Les corpus décrits ci-dessus comptent 454 occurrences de *prochain*, dont 290 à l'écrit et 164 à l'oral. Nous allons d'abord nous intéresser aux structures du type *N prochain*, c'est-à-dire quand l'adjectif est postposé au nom recteur.

3.1 Prochain postposé

Dans la partie orale, *prochain* au singulier (98 occurrences) ne s'emploie qu'avec des noms renvoyant à la notion de temps : *semaine, mois, an, année, lundi, vendredi, décembre, été, week-end, fois, siècle, trimestre, etc.* :

- (1) elle va la voir parce qu'ils arrivent la semaine prochaine [Oral, PFC]
- (2) et le grand plaisir est de le voir repris euh à l'opéra de Genève en décembre prochain [Oral, Nancy]

L'antéposition est ici impossible (**en prochain décembre*) ou alors provoque un changement de sens notable (*le week-end prochain* ≠ *le prochain week-end*). Au pluriel, nous n'avons trouvé aucun exemple où *prochain* est postposé.

Il existe également une contrainte sur le déterminant. Les noms de jours de la semaine et de mois s'en passent obligatoirement. Dans les autres cas, il faut un déterminant défini.

A l'écrit (122 occurrences), les noms recteurs sont majoritairement de type temporel (105 occurrences), mais il y a

10 autres noms (*mort, douceur, victoire, vote, etc.*) – presque toujours avec un déterminant défini – où *prochain* acquiert le sens de « proche » ou « imminent », que l'on relève essentiellement dans les textes littéraires. Ce sens-là ne se retrouve pas dans la partie orale.

Trois cas sont singuliers à plusieurs égards. Il s'agit de *genre prochain, station prochaine* et *rentrée prochaine* :

- (3) la spécificité formelle du syllogisme rhétorique par rapport au genre prochain du syllogisme dialectique n'est rien moins que clair dans Aristote [Ecrit, CNRS Edition]
- (4) j'entendais le sifflement des trains qui, plus ou moins éloigné, comme le chant d'un oiseau dans une forêt, relevant les distances, me décrivait l'étendue de la campagne déserte où le voyageur se hâte vers la station prochaine. [Ecrit, roman]
- (5) La mise à la disposition, à la rentrée prochaine, d'un important contingent supplémentaire d'emplois jeunes. [Ecrit, Jospin]

Le premier renvoie à une terminologie spécifique et prend le sens de « proche », le second se trouve dans une tranche bien définie, à savoir les romans antérieurs au XX^e siècle et le dernier peut librement se postposer ou s'antéposer sans modifier notablement le sens du syntagme².

Au pluriel, nous n'avons trouvé que 9 occurrences. Celles-ci sont majoritairement (5 occurrences) liées à la présence d'une coordination comme dans *les mois ou les années prochaines*. Dans ce contexte, l'antéposition est extrêmement difficile³. Une fois de plus, la tranche contenant les romans présentent des cas tout à fait spécifiques : *les cellules / intimités / vacances prochaines*. Dans les discours de J. Chirac et F. Mitterrand, on retrouve une fois l'expression *dans les mois prochains*.

2 Berthonneau note que cette « alternance entre postposition et antéposition est naturelle dans les situations qui impliquent la répétition d'une activité ou d'un processus » (p. 119).

3 Nous n'avons pas trouvé d'exemples tels que :
? *les prochains mois ou années*.

*Quand un corpus rencontre un adjectif du troisième type.
Etude distributionnelle de prochain*

Du point de vue sémantique, d'après Berthonneau (2002 : 116), « *Le N prochain* réfère à la première occurrence de N qui, dans une suite, se présente après t0 ».

3.2 Prochain antéposé

Pour *prochain N*, l'analyse des données ne révèle pas de grandes différences entre l'oral et l'écrit, contrairement à ce qui a été détaillé ci-dessus. On peut toutefois noter une plus grande diversité en ce qui concerne les noms recteurs à l'écrit. Au singulier, la présence de noms temporels est minoritaire, *fois* tirant son épingle du jeu à l'oral, où il représente près de la moitié des emplois. Nous n'avons trouvé aucun exemple du type **la prochaine semaine / année* (mais cf. ci-dessous).

Siècle se comporte de manière tout à fait originale. En effet, nous avons trouvé 6 occurrences à l'écrit de *prochain siècle*. Cela montre qu'avec *fois* il fait partie de ces rares noms temporels avec lesquels *prochain* (au singulier) peut librement se postposer ou s'antéposer. A noter encore deux cas intéressants, pour lesquels la contrainte d'antéposition est liée à la présence d'un déterminant possessif et/ou d'un complément du nom :

- (6) ma prochaine semaine de vacances je n'sais même pas quand la prendre [oral, CFPP]
- (7) Carine avait enfin trouvé l'endroit « idéal » pour leur prochain week-end [Ecrit, roman]

L'emploi d'un déterminant indéfini ou possessif est d'ailleurs peu fréquent.

- (8) le parti au pouvoir a toutefois annoncé un prochain retour au bon vieux temps de l'Etat providence. [Ecrit, *Le Monde*]
- (9) les gens fixent à voix basse la date de leur prochaine visite [Ecrit, *Courrier International*]

Au pluriel, la plupart des 79 occurrences sont formées avec des noms temporels, souvent combinés avec *au cours des*, *dans les*, *pour les* suivis d'un numéral.

- (10) Le débat sur l'avenir de l'Union ne peut se substituer à la responsabilité que nous avons dans les six prochains mois [Ecrit, *Nouvel Observateur*]

A noter qu'il n'y a qu'un seul exemple pour *chapitre* (*Il faut introduire une écriture dans les prochains chapitres*) et, comme signalé par Berthonneau (2002 : 119), aucun pour *page*.

3.3 Synthèse des résultats

Dans les deux tableaux ci-dessous, nous résumons les principaux résultats de notre étude (les pourcentages sont à interpréter par colonne).

Tableau 1. Fréquence des divers emplois de prochain dans notre corpus écrit

ECRIT (290 occ.)	
Singulier (217 occ.)	
Antéposé (95 occ.)	Postposé (122 occ.)
Noms temporels (20%)	Noms temporels (86,1 %)
Autres (80 %)	Sens de « proche » (9 %)
	Divers (4,9 %)
Pluriel (73 occ.)	
Antéposé (64 occ.)	Postposé (9 occ.)
Noms temporels (62,5 %)	
Autres (37,5 %)	

Tableau 2. Fréquence des divers emplois de prochain dans notre corpus oral

ORAL (164 occ.)	
Singulier (149 occ.)	
Antéposé (51 occ.)	Postposé (98 occ.)
Noms temporels (51 %) (<i>la prochaine fois</i> (25 occ./26))	Noms temporels (100 %) <i>année / an</i> (48 occ./98) <i>semaine</i> (23 occ./98)
Autres (49 %)	
Pluriel (15 occ.)	
Antéposé (15 occ.)	Postposé (0 occ.)

A l'oral, *prochain* est toujours postposé lorsqu'il accompagne un nom temporel au singulier, sauf avec *fois* (très

*Quand un corpus rencontre un adjectif du troisième type.
Etude distributionnelle de prochain*

majoritairement antéposé). Ainsi, *siècle*, *fois* et *rentrée* ne possèdent pas le fonctionnement général des autres noms temporels car *prochain* peut s'antéposer ou se postposer sans changement de sens notable. La présence d'un complément de nom représente un autre contexte où *prochain* va pouvoir s'antéposer : *ma prochaine semaine de vacances*⁴. Sans le complément de nom, ce dernier exemple semble très difficile à réaliser. A l'écrit, il y a un usage que l'on ne retrouve pas à l'oral, à savoir *prochain* au sens de « proche ».

Au niveau des patrons syntaxiques récurrents, on retrouve au pluriel *au cours des / dans les X prochaines années* et la coordination des noms recteurs lorsque *prochain* est postposé (*dans les mois ou les années prochaines*). Nous relevons deux exemples dans la tranche des discours politiques, de *dans les mois prochains*. Mais la construction semble peu naturelle. En effet, hors coordination, la postposition de *prochain* avec un nom recteur temporel au pluriel est extrêmement rare.

Nous avons également noté une contrainte supplémentaire : avec un déterminant possessif, *prochain* est bloqué en antéposition (uniquement pour les noms temporels).

(11) Leur prochain week-end / *Leur week-end prochain.

Ce qui confirme la valeur déictique de *prochain* postposé, lié, comme le signale Berthonneau (2002), à la sphère personnelle.

Avec les autres types de nom, la postposition semble possible comme le signale Berthonneau (2002), même si cela est extrêmement rare dans nos données :

(12) Ma mort prochaine

Selon elle, la présence d'un déterminant possessif dans (12) implique une proximité temporelle. Comme il s'agit d'un événement unique, la seule interprétation possible est celle de *proche*. L'antéposition est considérée par Berthonneau (2002 : 117) comme inacceptable⁵.

4 On a, dans ce cas, un « ensemble clos [...] qui ne définit pas un ordre dans une série » (Berthonneau, 2002 : 108).

5 Nous en avons trouvé tout de même un exemple chez Balzac :

Et avec un nom comme *semaine*, la contrainte est encore plus forte. Il y a obligation d'avoir un complément de nom. Dans ce cas, *prochain* signifie au plus proche de t0 (cf. Berthonneau 2002 : 107 et suivantes) :

- (13) *Leur semaine prochaine / ?Leur prochaine semaine / Leur prochaine semaine de vacances / *Leur semaine prochaine de vacances

Pour finir, la plupart des exemples s'écartant des tendances générales se retrouvent dans la tranche des romans.

4. Discussion

Dans cette partie, nous allons commencer par aborder la question des noms recteurs avec lesquels *prochain* présente un fonctionnement « atypique ». Puis nous ferons une rapide comparaison de nos résultats avec ceux de Berthonneau (2002).

4.1 Noms recteurs particuliers

Comme nous l'avons signalé ci-dessus, quelques noms recteurs temporels ou proches de la notion de temporalité diffèrent dans leur fonctionnement. Ce sont les noms tels *fois*, *saison*, *siècle*, *décennie*, *millénaire* ou *rentrée*.

Contrairement à *semaine* ou *année* (au singulier) qui se construisent exclusivement avec *prochaine* postposé, on trouve très majoritairement *prochaine fois* (avec *prochaine* antéposé donc) dans nos données, même si les deux possibilités sont attestées. Berthonneau (2002 : 118) explique cette singularité par le fait que « contrairement à la suite des lundis ou des semaines, il s'agit donc toujours d'un ensemble clos [et que ce nom est] le marqueur générique de l'occurrence d'un événement. A ce titre, il se situe à la charnière entre les noms temporels et les noms d'entités matérielles ». Ainsi, comme pour les noms désignant des entités matérielles, *prochain* se retrouve très majoritairement antéposé.

Mais la conscience de sa prochaine mort rendit pour un moment au jeune homme l'assurance d'une duchesse qui a deux amants [*La peau de chagrin*]

*Quand un corpus rencontre un adjectif du troisième type.
Etude distributionnelle de prochain*

De même pour *siècle* et *rentrée*, les fréquences dans nos données sont presque identiques pour l'antéposition et la postposition, avec une tendance légère à la postposition. Le faible nombre d'exemples ne permet pas de mener une étude approfondie sur ce point. Sémantiquement les deux positions semblent équivalentes. Tous nos exemples sont au singulier et sont accompagnés d'un déterminant défini. Berthonneau signale ces exemples et remarque une modification par un ordinal pour *siècle* (*au XXI^e siècle (maintenant) prochain*), ce que nous n'avons pas trouvé dans nos données. L'auteur note, par ailleurs, une nuance de sens entre *la prochaine rentrée* (impliquant la proximité de l'événement) et *la rentrée prochaine*. Sur ce dernier point, la valeur de proximité nous semble plutôt acquise lorsque *prochain* est postposé (avec d'autres noms que *rentrée*), même si les jugements ne sont pas très sûrs :

- (14) Réforme rendue possible par l'ouverture prochaine d'un nouveau pôle de l'art lyrique à la Bastille. [Ecrit, *Le Monde*]
- (15) Il a fait savoir, le mardi 27 janvier, qu'il « rendra publique, dans les jours qui viennent, et déposera à la prochaine ouverture du Parlement, une nouvelle proposition de loi, qui reprendra le texte de la loi annulée et la complétera, pour prendre en compte les remarques du Conseil constitutionnel ». [Ecrit, *Le Monde*]⁶

En ce qui concerne *saison*, il n'y a pas de préférence pour une position particulière, mais une nuance de sens et une distribution contrôlée par le sujet traité dans le texte. *La prochaine saison* s'emploie au sujet des séries télévisées pour désigner l'intégralité des épisodes correspondant à l'année suivante et *la saison prochaine* est utilisée dans le registre sportif pour désigner la période d'activités sportives qui s'annonce.

⁶ Comme nous l'a très justement fait remarquer notre relecteur, les deux occurrences de *ouverture* ne sont peut-être pas similaires. En (14) on pourrait y voir un événement ponctuel et en (15) un événement cyclique. Or, nous ne savons pas dans quelle mesure ce paramètre influe sur la valeur de proximité prise par *prochain*.

- (16) Pour les fans de Buffy, sachez que tout va bien pour la série, les producteurs sont en train de négocier sur quelle chaîne américaine la prochaine saison sera diffusée. [Ecrit, *People*]
- (17) Pour la saison prochaine, le club recrute quelques filles désirant évoluer en championnat régional... [Ecrit, *Journal de Saône-et-Loire*]

Pour compléter la liste des noms recteurs temporels présents dans nos corpus, on peut signaler *le trimestre prochain* (1 occurrence à l'oral) ainsi que *le soir prochain* (1 occurrence dans un roman) et *les prochaines décennies* (3 occurrences à l'écrit). Par manque d'exemples, nous ne faisons que les nommer.

Par ailleurs, nous n'avons pas trouvé d'exemples avec *moment*, *instant*, *époque* ou *période*. Comme le précise Berthonneau (2002 : 110), cela peut se comprendre aisément car ils n'ont ni de structure séquentielle ni de statut d'individu, ce qui bloque la combinaison avec *prochain*.

4.2 Comparaison introspection / corpus

Nos conclusions et notre description ne sont pas fondamentalement éloignées de celles de Berthonneau (2002) qui pourtant a basé la plupart de ses observations sur des exemples construits. Nous avons ainsi pu vérifier la plupart des propriétés décrites par Berthonneau. Cela montre donc que la linguistique introspective et la linguistique sur corpus peuvent converger. Cependant, il existe un certain nombre de points sur lesquels nos observations divergent.

Tout d'abord, nous avons mis l'accent sur l'importance que revêt le pluriel pour l'étude de *prochain*, et plus particulièrement pour ses possibilités de placement. Ensuite, les études sur corpus permettent de mettre en évidence des tendances et des fonctionnements largement majoritaires, qui demeurent inaccessibles autrement. Par exemple, l'antéposition de *prochain* à un nom temporel est extrêmement rare (hormis pour *fois* qui représente un cas atypique). Il y a donc une très forte tendance à la postposition. Pour *fois*, si *prochain* peut effectivement se postposer ou s'antéposer sans modification fondamentale de sens, la tendance est très nettement à l'antéposition. De même, les cas de postposition à des noms autres que temporels sont totalement

*Quand un corpus rencontre un adjectif du troisième type.
Etude distributionnelle de prochain*

marginaux et même inexistants dans la partie orale. Ce constat contredit l'argumentation de Berthonneau sur le fait que *prochain* pourrait s'appliquer à des entités matérielles à condition que le N s'inscrive dans une suite prédictible. Or cette argumentation s'appuie sur des exemples dont on peut interroger l'acceptabilité. Il en est ainsi de la série proposée dans ce qui suit (*femme / copine / maire*) et qui oppose des entités selon qu'elles peuvent ou non présupposer « une structure séquentielle a priori ». Sur ce point, l'écart entre données observées et données construites est très net.

- (18) *Marie, c'est la femme prochaine de Paul.
- (19) Il aime les grandes femmes. Tu verras, sa copine prochaine, elle fera encore dix centimètres de plus que lui.
- (20) ?*Le maire prochain est blond.
- (21) Le maire prochain devra mettre un terme à ces pratiques.
- (22) *Le berceau servira à votre enfant prochain.

Nous n'avons trouvé aucune attestation de N animé + *prochain* dans nos corpus. En étendant la recherche à *Google-Recherche de Livres*, nous en avons relevées quelques-unes dans des ouvrages modernes (littérature, essais, revues), mais en nombre insuffisant pour pouvoir mener une étude systématique. Ainsi, la nuance de sens que Berthonneau essaie de mettre en évidence est fragilisée par l'absence de cas similaires dans nos données et des jugements somme toute assez ténus. Nous avons le plus grand mal à partager ses intuitions.

L'analyse de Berthonneau pose ainsi le problème de la représentativité des données obtenues par introspection. Quel que soit le jugement que l'on puisse porter sur ces derniers exemples, force est de constater que nous n'avons relevé aucune attestation de ce type dans nos corpus. Il en est de même pour les entités *chapitre*, *page*, *minute*, *seconde*, etc. pour lesquels nous n'avons relevé aucun exemple.

De même, Berthonneau accepte *Le (prochain) chapitre (prochain)* et rejette **la (prochaine) page (prochaine)* et **?A la prochaine (minute / seconde), on y verra plus clair*. Dans nos données, nous n'avons d'exemples ni de l'une ni de l'autre

tournure. Et en élargissant notre recherche, nous trouvons des exemples considérés par Berthonneau comme non acceptables :

- (23) Ne PAS commencer une nouvelle section directement en bas de la page, mais transférer le titre jusqu'au début de la prochaine page. [www.cinam.univ-mrs.fr/cristal6/template/Cristal6.doc]⁷
- (24) On fait l'hypothèse que l'appareil ne vieillit pas : la probabilité qu'il tombe en panne pour la première fois dans la prochaine seconde (ou la prochaine minute, ou le prochain jour) ne dépend pas de l'âge de l'appareil (c'est-à-dire du temps pendant lequel l'appareil n'est jamais tombé en panne). [extrait d'un cours de statistique : http://uuu.enseirb-matmeca.fr/~duchon/ProbaStat/Notes_IS101.pdf]

Sur ce dernier exemple, après avoir rejeté les emplois du type *prochaine minute / seconde*, elle en cite deux pour lesquels la position sujet (« conférant au référent une certaine individualité ») ou le fait que « le segment de temps [soit] considéré dans son étendue » (p. 111, note 14) les rendent acceptables. Berthonneau rejette donc dans un premier temps ce qu'elle trouve possible par la suite dans certaines conditions, et ce, à quelques lignes d'intervalles.

Les jugements qu'elle utilise pour distinguer les chapitres et les pages d'une part et les secondes / minutes / heures d'autre part doivent donc être nuancés. Et il en va de même pour l'analyse qu'elle propose de la tournure *station prochaine* qui se trouve au centre des principales conclusions de l'étude.

Il semble donc que, sur la question des jugements de grammaticalité, les deux approches comparées ci-dessus soient irréductibles l'une à l'autre. L'ensemble des exemples précédents montre donc l'intérêt de confronter les données attestées à celles obtenues par introspection, bien que la visée de Berthonneau ne soit pas une analyse distributionnelle des items, mais plutôt une réflexion de leur valeur en fonction des dimensions du temps et de l'espace.

⁷ On peut toutefois noter ici que la page dont il s'agit n'est pas une page de texte mais une page sur un écran. Il est alors possible qu'elle soit « individualisée », ce qui pourrait expliquer cet emploi.

5. Comparaison avec des entrées de dictionnaires

Il nous a semblé intéressant de poursuivre notre réflexion en comparant notre description à celles des dictionnaires. On peut en effet se demander dans quelle mesure ceux-ci reflètent les usages authentiques ainsi que les grandes tendances répertoriées ci-dessus, étant donné que leurs définitions se basent essentiellement sur des extraits littéraires ou des exemples construits. Nous avons fait le choix d'examiner trois ouvrages, le *Trésor de la Langue Française* (TLF), le *Grand Larousse de la Langue Française* (GLLF) et le *Petit Robert* édition 2000 (PR).

Ces trois dictionnaires ont en commun d'ordonner de la même manière les acceptions de l'adjectif *prochain*. Par exemple, la première acception du lexème est le sens locatif, alors que nous avons pu constater que nous trouvons bien moins la valeur locative de *prochain* que la valeur temporelle, et cela tant dans les corpus oraux que dans les corpus écrits. Nous pouvons relever cependant l'effort du TLF et du PR pour mettre l'accent sur la position de l'adjectif par rapport au nom recteur alors que le GLLF n'en fait pas mention. Le TLF mentionne fréquemment le critère de position de l'adjectif comme étant un facteur déclencheur de sens. Il indique, en effet, que l'adjectif *prochain* est postposé lorsqu'il « qualifie un événement personnel, unique ou considéré comme tel » comme c'est le cas dans *victoire prochaine*. De même, le TLF et le PR insistent sur le fait que l'adjectif doit être antéposé quand ce dernier implique une notion de série : « (Antéposé) Qui viendra immédiatement après (dans une série) » (*Le Petit Robert*, 2000).

Après avoir mis en valeur les utilisations locatives et temporelles de *prochain*, les trois dictionnaires évoquent les combinatoires à la fin de chaque article. Parmi elles, sont communes aux trois ouvrages : *genre prochain*, *pouvoir prochain*, *effet / cause prochain(e)*. Dans nos corpus, seul *genre prochain* apparaît, et ce uniquement à l'écrit et à hauteur d'une seule occurrence.

Nous avons relevé les noms recteurs de l'adjectif *prochain* les plus fréquemment utilisés et avons pu constater que *année*, *an* et *semaine* étaient les plus représentés. Chacun des trois dictionnaires utilise au moins un de ces noms en guise

d'exemple d'une réalisation de *prochain*, ce qui montre l'intégration de cette combinatoire dans le lexique.

En revanche, nous n'avons relevé qu'une seule occurrence du nom recteur *station* (dans la partie roman). Berthonneau (2002 : 106) insiste pourtant sur ce cas, le considérant comme étant une représentation locative où « la station qui suit [est] celle où l'on est ou que l'on vient de quitter ». Si le GLLF ne présente pas le nom *station* dans les exemples qu'il donne, le TLF cite bien *la station prochaine* (en indiquant toutefois que l'adjectif est généralement antéposé avec *station*) de même que le PR. Ce dernier ne mentionne pas la place de l'adjectif par rapport au nom.

Le décalage entre l'absence dans nos corpus de la cooccurrence *station prochaine* et sa mention dans deux dictionnaires sur trois illustrent la difficulté que pose cet exemple. En effet, en tant que locuteur, nous percevons qu'il existe une nuance entre *prochaine station* et *station prochaine* (pour peu que l'on accepte ces deux tournures, ce qui n'est déjà pas incontestable). Cependant, nous n'identifions pas exactement quel changement de sens s'opère entre elles. Sur ce point précis, les dictionnaires ne nous sont d'aucun secours. En effet, le TLF mentionne que l'adjectif est généralement antéposé et donne un exemple où l'adjectif est postposé au nom. De son côté, le PR cite seulement un exemple dans lequel *prochain* est antéposé au nom *station*.

Nous avons vu précédemment que, parmi les expressions retenues par les trois dictionnaires (*genre prochain*, *pouvoir prochain*, *effet / cause prochain(e)*), seule *genre prochain* apparaît dans nos corpus (et encore, à une seule reprise). Cela laisse à penser que ces combinaisons sont soit très peu usitées, soit utilisées dans des contextes particuliers qui ne sont pas représentés dans nos corpus. Il est par ailleurs difficile de savoir quels critères ont guidé la sélection des noms recteurs à prendre en compte avec *prochain* dans les dictionnaires. Néanmoins, le dictionnaire oriente le plus souvent la description sur le lexique de la langue sans prendre en compte l'utilisation du lexique et les habitudes conversationnelles. Ainsi, la fréquence d'utilisation d'une expression ne sera pas un critère pertinent pour sa

*Quand un corpus rencontre un adjectif du troisième type.
Etude distributionnelle de prochain*

description. S'il existe de nombreux dictionnaires savants et commerciaux qui tentent d'être exhaustifs dans leur description, les études faites sur l'oral montrent le besoin d'une nouvelle forme de dictionnaires dans lesquels les usages seraient un des critères de classification dans chaque entrée, à l'image du *Collins Cobuild English Dictionary* pour l'anglais.

6. Conclusion

Pour le français, il existe relativement peu d'études distributionnelles réalisées à partir de données aussi diversifiées que celles utilisées dans cette étude et en particulier pour des adjectifs marginalisés comme *prochain*. L'étude sur corpus permet néanmoins de mettre en évidence les divers fonctionnements de l'adjectif *prochain* et de comparer ces usages aux analyses recourant à des données majoritairement introspectives.

Nous avons ainsi pu montrer que la description de *prochain* dépend non seulement du nom qui l'accompagne, mais également de la position relative Nrecteur-*prochain*, du nombre, du déterminant et d'éventuels modificateurs. Les sélections lexicales sont d'ailleurs largement conditionnées par la position de l'adjectif, ce que l'oral non planifié permet de montrer avec encore plus de clarté que l'écrit élaboré. Nous avons aussi observé quelques noms recteurs temporels (*fois*, *siècle*, etc.) qui ne possèdent pas le même fonctionnement que les emplois les plus représentés (*semaine*, *année*, etc.). Notre démarche nous a permis de relever des phénomènes très ciblés, à l'image de *saison* qui, accompagné de *prochain*, présente une corrélation entre sa position, un sens spécifique et le sujet traité. En effet, il y a une différenciation entre *saison prochaine* dans le registre sportif et *prochaine saison* pour les séries télévisées. Nous avons donc pu aborder aussi bien les grandes tendances que les phénomènes plus difficilement saisissables.

Par ailleurs, nous avons également mis en évidence les écarts qui pouvaient exister entre notre présentation basée en grande partie sur la fréquence et celle des dictionnaires qui prennent peu (voire pas du tout) en compte ce paramètre. C'est particulièrement visible pour l'emploi locatif quasi-inexistant dans nos données et considéré comme le premier sens dans les

trois dictionnaires pris en compte. De même, Berthonneau (2002) met l'accent sur des cas qui, dans nos corpus, sont soit inexistants, soit marginaux, soit liés à des textes appartenant au registre romanesque. En fait, l'auteur ne considère pas la fréquence comme pertinente. Alors qu'elle refuse les exemples *à la minute prochaine, la bombe va exploser et ?*A la prochaine minute, on y verra plus clair, elle reconnaît quelques pages plus bas qu'il y a une occurrence de *prochaine minute* (en tant que sujet) et deux de *minute prochaine* dans le corpus qu'elle consulte (Frantext 1950-2000). Du coup, quand il s'agit d'exemples très peu fréquents, on peut légitimement se demander pourquoi certains sont considérés comme agrammaticaux (à l'image de *prochaine minute*) et d'autres à la base de sa démonstration (à l'image de *station prochaine*) ? Certains exemples centraux pour son analyse se révèlent donc très peu attestés et posent le problème des limites de l'intuition du linguiste et de la place prépondérante qu'occupe le français littéraire par rapport au français en général.

Au-delà de l'analyse de détail, le présent travail doit être considéré comme une étape dans l'élaboration d'une base descriptive solide permettant de problématiser le statut catégoriel de *prochain*. Nous avons commencé à étudier d'autres adjectifs, tels *dernier* et *suivant*, dont les fonctionnements sont similaires à *prochain* sans pour autant être identiques. De ce travail en cours, il ressort qu'il est nécessaire, comme le signale Schnedecker (2002a, b), de « procéder à l'analyse ponctuelle et systématique des unités à classer ». Nous espérons avoir montré que l'analyse sur corpus diversifiés peut largement contribuer à l'avancée de cette problématique et à une meilleure connaissance de la classe foisonnante des adjectifs.

Références bibliographiques

- Berthonneau A.-M. (2002). « *Prochain / dernier* et compagnie. Les adjectifs “déictiques” à l'épreuve de l'espace ou comment circuler dans le temps, l'espace, le texte », *Langue Française* 136, 1 : 104-125.
- Blasco-Dulbecco M. & Cappeau P. (2004). « Quelques remarques sur l'adjectif à l'oral », in J. François (dir.),

*Quand un corpus rencontre un adjectif du troisième type.
Etude distributionnelle de prochain*

L'adjectif en français et à travers les langues. Caen : Presses Universitaires de Caen (Bibliothèque de Syntaxe & sémantique), 413-428.

- Blasco-Dulbecco M. & Cappeau P. (2005). « Ce que les corpus oraux nous apprennent sur les adjectifs », in G. Williams (dir.), *La linguistique de corpus*. Rennes : Presses Universitaires de Rennes, 69-80.
- Branca-Rosoff S., Fleury S., Lefevre F. & Pires M. (2009). *Discours sur la ville. Corpus de Français Parlé Parisien des années 2000.* (CFPP2000).
- Delic (2004). « Présentation du Corpus de Référence du Français Parlé », *Recherches sur le français parlé* 18 : 11-42.
- Durand J., Laks B. & Lyche C. (2002). « La phonologie du français contemporain : usages, variétés et structure », in C. Pusch & W. Raible (éd.), *Romanistische Korpuslinguistik – Korpora und gesprochene Sprache / Romance Corpus Linguistics – Corpora and Spoken Language*. Tübingen : Gunter Narr Verlag, 93-106.
- Durand J., Laks B. & Lyche C. (2005). « Un corpus numérisé pour la phonologie du français », in G. Williams (éd.), *La linguistique de corpus*. Rennes : Presses Universitaires de Rennes, 205-217.
- Frantext = INaLF, puis ATILF (1992). *Frantext (outil de consultation de ressources informatisées sur la langue française)*, Nancy, CNRS/ATILF, site internet : <http://www.frantext.fr>.
- Goes J. (1999). *L'adjectif entre nom et verbe*. Louvain : Duculot.
- Google, Recherche de Livres = 2007-. Google, moteur de recherche, site internet : <http://www.google.fr/books>.
- Grand Larousse de la langue française en six volumes = Guilbert L., Lagane R., Niobey G., avec le concours de Bonnard H., Casati L., Lerond A. *et al.* Paris : Larousse ; t. 1 (A - Cip), 1971 ; t. 2 (Cir - Ery), 1972 ; t. 3 (Es - Inc), 1973.
- Noailly M. (1999). *L'Adjectif en français*. Paris / Gap : Ophrys.
- Robert = Rey-Debove J. & Rey A. (2000). *Dictionnaire alphabétique et analogique de la langue française*. Paris : Dictionnaires Le Robert.

- Schnedecker C. (2002a). « Présentation : les adjectifs “inclassables”, des adjectifs du troisième type ? ». *Langue Française* 136, 1 : 3-19.
- Schnedecker C. (2002b). « Premier, principal, primordial : des adjectifs qui sortent du rang ? ». *Langue Française* 136, 1 : 89-103.
- Sinclair J. (1997). *Collins Cobuild English Dictionary*. London : Harper Collins.
- TLF = Imbs P. & Quemada. B. (dir.), (1971-1994). *Trésor de la Langue Française*. Dictionnaire de la langue du XIX^e et du XX^e siècle (1789-1960), 16 vol. Paris : Editions du CNRS / Gallimard.