



HAL
open science

Traitement lexicographique des collocations à collocatif actanciel

Alain Polguère

► **To cite this version:**

Alain Polguère. Traitement lexicographique des collocations à collocatif actanciel. Inès Sfar; Pierre-André Buvet. La phraséologie entre fixité et congruence. Hommage à Salah Mejri, Academia-L'Harmattan, pp.265-287, 2018, 978-2-8061-0334-5. hal-03071962

HAL Id: hal-03071962

<https://hal.science/hal-03071962>

Submitted on 16 Dec 2020

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.

Traitement lexicographique des collocations à collocatif actanciel*

Alain Polguère
Université de Lorraine, CNRS, ATILF

1. Introduction

Selon la vision fonctionnelle, non statistique, de la phraséologie (Hausmann 1989 ; Mel'čuk 1995), la collocation est un cas particulier de phrasème qui peut être caractérisé au moyen de la définition suivante, sans doute imparfaite, mais suffisante pour entamer la présente discussion¹.

Une collocation prototypique est un syntagme construit A + B (*le travail_A s'amoncelle_B*) ou B + A (*crouler_B sous le travail_A*) :

- où la lexie A – la base de la collocation – est sélectionnée par le locuteur en fonction de ses caractéristiques individuelles
- et où B – le collocatif de A dans la collocation – est sélectionné en premier lieu en fonction des propriétés individuelles de combinatoire de A pour exprimer un contenu sémantico-syntaxique donné auprès de A.

La description linguistique des collocations relève fort logiquement de la description lexicographique de leur base, puisque le fait de contrôler un collocatif donné est une caractéristique de combinatoire lexicale de la base. Dans le cadre de la Lexicographie Explicative et Combinatoire ou LEC (Mel'čuk et coll. 1995), une telle caractéristique se modélise au moyen du système des fonctions lexicales syntagmatiques (Mel'čuk 1996). Il existe cependant en langue de très nombreuses collocations – du type *hurlement de plaisir* ou *cri d'orgasme* – où le collocatif (*plaisir*, *orgasme*) peut être qualifié d'« actanciel », car il participe à l'expression d'un des actants de la base prédicative (*hurlement*, *cri*). La description lexicographique de telles collocations est doublement problématique. Tout d'abord, elle relève autant de la description du régime syntaxique de la base que de celle de sa combinatoire lexicale restreinte. Ensuite, ces collocations peuvent généralement recevoir une analyse fonctionnelle inverse, où l'élément syntaxiquement régi peut être conçu comme base et le régissant comme collocatif : *hurlement de plaisir* serait ainsi tout autant une collocation de *hurlement* qu'une collocation de *plaisir*.

Nous commencerons notre exposé par une brève présentation de la modélisation formelle de la combinatoire restreinte syntaxique et lexicale des lexies, telle qu'effectuée dans le cadre de la LEC (section 2). Nous définirons ensuite ce qu'il faut entendre par collocatifs actanciels et montrerons en quoi il s'agit d'un phénomène omniprésent dans les lexiques des langues (section 3). Finalement (section 4), nous expliquerons comment, selon nous, il convient de rendre compte lexicographiquement des collocations à collocatifs actanciels. Il ne nous sera pas possible d'introduire de façon détaillée toutes les notions lexicologiques et lexicographiques sur lesquelles

* Dans Inès Sfar et Pierre-André Buvet (dir.) : *La phraséologie entre fixité et congruence. Hommage à Salah Mejri*. Coll. « Sciences du langage : Carrefours et points de vue » 19, Louvain-la-Neuve, Academia-L'Harmattan, pp. 265–287, 2018.

1 Nous remercions chaleureusement Dorota Sikora pour ses commentaires sur une version préliminaire du présent article.

repose notre analyse. Nous nous limiterons, dans de nombreux cas, à une série de renvois bibliographiques qui, nous l'espérons, permettront au lecteur non familiarisé avec la LEC et, plus généralement, avec l'approche de la théorie linguistique Sens-Texte, de trouver un complément d'information. Cela étant dit, il nous semble que les notions auxquelles nous faisons appel relèvent, pour l'essentiel, d'un savoir de plus en plus partagé par les chercheurs impliqués dans l'étude et la modélisation du lexique.

2. Combinatoires restreintes syntaxique et lexicale

La LEC postule un type particulier de modèle lexicographique, appelé *Dictionnaire Explicatif et Combinatoire* (Mel'čuk et coll. 1984, 1988, 1992, 1999). Dans un tel modèle, la description des collocations contrôlées par la lexie-vedette d'un article lexicographique relève, plus généralement, de la description de la combinatoire dite *restreinte* de la lexie. Il s'agit de la combinatoire de la lexie qui ne peut être déduite ni de son sens (décrit par une définition lexicographique) ni de sa forme. Pour les besoins du présent exposé, nous allons nous concentrer ici sur deux éléments de la description de la combinatoire restreinte des lexies : la combinatoire syntaxique ou régime de la lexie-vedette d'un article lexicographique et sa combinatoire lexicale – qui correspond à l'ensemble des collocations dont la lexie-vedette est la base.

2.1 Régime syntaxique de la lexie-vedette d'un article lexicographique

Il est maintenant bien admis que toute unité lexicale de la langue ayant une valeur prédicative – qu'il s'agisse d'un verbe, d'un nom, etc. – tend à régir des structures syntaxiques particulières dans le cadre de l'expression des actants sémantiques qu'elle contrôle. La prise en compte de cette propriété des lexies, leur *valence syntaxique active* ou *régime syntaxique*, est de plus en plus présente dans les dictionnaires de langue et les bases de données lexicales. Elle est au cœur de la modélisation lexicale dans des approches aussi diverses que le lexique-grammaire (Gross 1975), la sémantique des cadres angl. *frame semantics* (Fillmore 1982 ; Baker et coll. 2003), l'approche pronominale (Van den Eynde & Miertens 2003) ou les classes d'objets (Gross 2008) ; elle joue bien entendu un rôle crucial dans la LEC, qui nous concerne tout particulièrement ici. Alors que la notion de régime syntaxique était traditionnellement associée avant tout à la partie du discours des verbes, sa pertinence pour les lexies prédicatives des autres parties du discours est maintenant bien reconnue, notamment pour ce qui est de la description lexicographique des unités nominales prédicatives (Herbst 1988 ; Herbst et coll. 2004 ; Van den Eynde 1998).

Du point de vue lexicographique, il existe de multiples techniques de prise en compte de la valence lexicale. Nous les présentons maintenant très brièvement, en commençant par les techniques les moins explicites et formelles.

Le *Grand Dictionnaire Encyclopédique Larousse* (GDEL 1982–1985) – dictionnaire dont la carrière commerciale a malheureusement été éphémère – a joué un rôle de pionnier en lexicographie française en faisant systématiquement précéder la définition de chaque sens verbal par un patron de complémentation en italique, qui vise à décrire de façon simple et immédiatement compréhensible par le non-spécialiste le régime de la lexie en question². Cela est illustré ci-dessous par l'entrée ENSEIGNER du GDEL.

2 À notre connaissance (mais, nous ne sommes nullement métalexico-grahe), le GDEL a non seulement été pionnier dans ce domaine, en lexicographie française commerciale, mais il est aussi demeuré le seul à décrire systématiquement le régime dans ses articles. Cette description – introduite sous l'impulsion de Jean Dubois – ne se retrouve, nous semble-t-il, dans aucun autre dictionnaire commercial français.

ENSEIGNER v. t. (lat. pop. *insignare*, du lat. class. *insignire*, signaler) [conj. 4].
1. Enseigner une science, un art, une discipline à qqn, à un groupe, le lui faire apprendre, le lui expliquer en lui donnant des cours, des leçons : *Enseigner les mathématiques à de jeunes enfants.* — **2.** Enseigner (qqpart), sans compl. d'objet, (y) être enseignant : *Il enseigne à l'université.* — **3.** Enseigner à qqn qqch (abstrait), que + ind., enseigner à + inf., enseigner + interr. indirecte, le lui faire savoir, le lui faire connaître, le lui inculquer par une sorte de leçon, en parlant de qqn ou de qqch ; apprendre, montrer (langue soutenue) : *Elle m'avait enseigné les usages de ce monde, comment vivre en société. Enseigner un moyen, une méthode. Enseigner à se passer des autres. La vie se chargera de le lui enseigner. L'histoire nous enseigne que tout est recommencement.* — **4.** Litt. Enseigner qqn, le guider dans ses apprentissages, être son maître, son professeur ; instruire : *Enseigner des adolescents.* — **5.** Litt. Enseigner qqch (concret) à qqn, le lui indiquer, le lui signaler (vieilli) : *Enseigner son chemin à un passant, un hôtel à un ami.*
 ◆ **s'enseigner** v. pr., être enseigné v. pass. Faire l'objet d'un enseignement : *Le russe ne s'enseigne pas, n'est pas enseigné dans tous les collèges.*

Figure 1 : Entrée ENSEIGNER dans le GDEL (1982–1985)

Voilà comment le GDEL présente dans son introduction les formules employées pour décrire les régimes syntaxiques :

« Les sens des verbes, et éventuellement de certains adjectifs, sont distingués selon leurs constructions syntaxiques ; celles-ci, indiquées pour chaque sens, comportent des informations sur la nature et la forme des compléments. Le caractère facultatif des compléments d'un verbe est indiqué par la présence de parenthèses. Pour les autres catégories, la définition n'est précédée de l'indication de la construction que lorsque celle-ci détermine clairement une distinction de sens. » (GDEL 1982–1985 : vol. 1, VI)

Il est intéressant de noter que c'est presque en contrebande que le régime est introduit dans les articles du GDEL, puisque sa description est présentée comme une information permettant de distinguer les différents sens d'un vocable. Il n'est nullement fait mention dans la courte introduction du GDEL du fait que l'information sur le régime peut aussi, et surtout, aider l'utilisateur du dictionnaire à apprendre quelles structures syntaxiques sont contrôlées par une lexie donnée et donc, à apprendre comment employer cette lexie correctement dans une phrase.

C'est avec une approche au contraire résolument pédagogique et orientée vers la production linguistique que procède le dictionnaire commercial grand public qui a introduit avec le plus de systématisme (et avec le plus de succès) le régime de ses lexies-vedettes : le *Collins Cobuild dictionary* (Sinclair 1987). Les indications grammaticales données dans les marges, en regard des articles correspondants, contiennent notamment la description des structures syntaxiques régies sous forme de formules compactes et relativement transparentes. Cette extraction en marge de ce type d'information, dans le cadre d'un dictionnaire commercial, nous semble d'ailleurs une idée brillante. La version en ligne du dictionnaire tire avantage de l'absence de la contrainte d'espace, qui nuit tant aux dictionnaires papier, pour introduire les structures régies non plus simplement par une énumération de formules codées (comme dans la version papier), mais en accompagnant chacune de ces formules d'un court exemple lexicographique illustrant le régime en question. Cela est illustré par l'entrée du verbe TEACH 'enseigner' ci-dessous.

teach (**teaches** 3rd person present) (**teaching** present participle) (**taught** past tense & past participle)

1 **verb** If you teach someone something, you give them instructions so that they know about it or how to do it.
 The trainers have a programme to teach them vocational skills... V n n
 George had taught him how to ride a horse... V n wh
 She taught Julie to read... V n to-inf
 The computer has simplified the difficult task of teaching reading to the deaf. V n to n

2 **verb** To teach someone something means to make them think, feel, or act in a new or different way.
 Their daughter's death had taught him humility... V n n
 He taught his followers that they could all be members of the kingdom of God... V n that
 Teach them to voice their feelings. V n to-inf, Also V n wh, V n about n

3 **verb** If you teach or teach a subject, you help students to learn about it by explaining it or showing them how to do it, usually as a job at a school, college, or university.
 Ingrid is currently teaching Mathematics at Shimla Public School... V n
 She taught English to Japanese business people... V n to n
 She has taught for 34 years... V
 She taught children French. V n n
 ...this twelve month taught course. V-ed

4
 → teaching
 → to teach someone a lesson
 → lesson

Figure 2 : Entrée du verbe TEACH ‘enseigner’ dans le *Collins Cobuild dictionary* en ligne³

Il existe bien entendu de multiples façons de décrire les régimes syntaxiques dans les ouvrages de linguistique, les bases lexicales spécialisées – comme le *Dicovalence* (Van den Eynde & Miertens 2006) – et les dictionnaires théoriques du type *Dictionnaire explicatif et combinatoire* (Mel’čuk et coll. 1984, 1988, 1992, 1999). Nous nous limiterons ici à ce dernier, puisqu’il relève de la LEC, qui est notre approche théorique et descriptive de référence.

Le dictionnaire postulé par la LEC présente le régime syntaxique des lexies-vedettes sous forme d’un tableau, appelé *tableau de régime* ou aussi parfois *schéma de régime*, mettant en correspondance les structures actanciennes sémantique, syntaxique profonde et syntaxique de surface des lexies. La théorie Sens-Texte, dont la LEC est la composante lexicale, s’appuie en effet sur trois niveaux de structure actancielle, qu’il ne nous appartient pas de décrire ici. On pourra se reporter à Mel’čuk (2004a,b) pour la présentation détaillée de la théorie des actants élaborée dans l’approche Sens-Texte. Quant aux tableaux de régime proprement dits, leur principe est décrit dans, notamment, Milićević (2009) et nous nous contentons de donner ci-dessous, à titre d’illustration, la définition et le tableau de régime de la lexie-vedette ENSEIGNER **2b** dans le premier volume du *Dictionnaire explicatif et combinatoire du français contemporain*, désormais DECFC⁴.

3 <http://dictionary.reverso.net/english-cobuild/teach> [page consultée le 12 octobre 2012]

4 Nous avons sélectionné cette acception du vocable ENSEIGNER, plutôt que son acception de base, uniquement parce que son tableau de régime est relativement compact et facile à inclure ici.

2b. *X* enseigne *Y* à *Z* = *X* contient une affirmation Y_1 qui fait partie d'une doctrine Y_2 proposée 4a dans *X* [comme si *X* enseignait 2a *Y* à *Z*].

Régime

1 = X	2 = Y	3 = Z
1. N	1. N 2. <i>que</i> PROP obligatoire	1. à N

$C_1 + C_2$: *La Bible enseigne la transcendance de Dieu <que Dieu est transcendant>*

$C_1 + C_2 + C_3$: *La Bible nous enseigne que Dieu est transcendant*

Figure 3 : Définition et tableau de régime de ENSEIGNER 2a dans Mel'čuk et coll. (1984 : 96)

Comme on le voit, le tableau de régime a pour finalité d'associer à chaque position actancielle sémantique contrôlée par la lexie-vedette l'ensemble des structures syntaxiques qui en supportent l'expression. On notera que, dans les faits, les tableaux de régime du type ci-dessus sont des présentations simplifiées. Ils devraient mentionner explicitement le nom des dépendances syntaxiques de surface⁵ qui lient les lexies-vedettes à leurs dépendants régis⁶.

2.2 Collocations contrôlées par la lexie-vedette

Nous serons beaucoup plus bref sur le sujet de la modélisation lexicographique des collocations dans les dictionnaires commerciaux et les ressources lexicales modernes. Il existe de très nombreuses publications sur le sujet – par exemple, Hausmann (1989), Moon (2008) ; nous nous contenterons donc de rappeler la façon dont la LEC traite la question.

Dans l'approche de la LEC, les collocations types – c'est-à-dire, celles qui relèvent de patrons récurrents en langue – sont décrites au moyen du système des *fonctions lexicales standard* (Mel'čuk 1996). Plus précisément, le système des fonctions lexicales standard peut être subdivisé en *fonctions lexicales paradigmatisques* – qui rendent compte des liens paradigmatisques de dérivation sémantique – et *fonctions lexicales syntagmatiques* – qui rendent compte des liens syntagmatiques base-collocatif. Un même type d'outil notionnel et formel – les fonctions lexicales standard – permet donc de rendre compte à la fois des liens paradigmatisques et syntagmatiques qu'une lexie-vedette entretient avec d'autres lexies de la langue. Les articles lexicographiques de la LEC contiennent ainsi une zone de fonctions lexicales, dont une sous-partie relève spécifiquement de la description des collocations, décrites au moyen des fonctions lexicales syntagmatiques. Dans le cadre du projet de construction de la base lexicale DiCo (Mel'čuk & Polguère 2006), un double encodage des liens lexicaux (paradigmatiques et syntagmatiques) a été introduit, puisqu'à l'encodage formel par fonctions lexicales s'est ajouté un encodage « grand public » effectué au moyen de *formules de vulgarisation* paraphrasant les liens lexicaux en question. Cette approche est reprise dans le cadre du projet de construction du Réseau Lexical du Français ou RLF (Lux-Pogodalla & Polguère 2009) et nous l'illustrons ci-dessous avec un extrait de la zone de fonctions lexicales de l'article AMIN I.1 du RLF dans son état présent⁷ : il s'agit d'une partie des liens

5 Sur la notion de dépendance syntaxique de surface, voir Iordanskaja & Mel'čuk (2009).

6 Cf. les tableaux de régime plus explicites présentés dans Milićević (2009).

7 La construction du RLF étant en cours de réalisation, les données que nous sommes en mesure de présenter sont nécessairement appelées à évoluer.

syntagmatiques contrôlés par la lexie-vedette. Dans l’affichage ci-dessous, les formules de vulgarisations précèdent immédiatement l’encodage par fonctions lexicales ; les symboles \$1 et \$2 dénotent, respectivement, le premier et le second actant de la lexie-vedette (‘individu X, ami I.1 de l’individu Y’).

Avec qui la relation est forte
Magn **bon**_{Adj} IV | antépos, **grand**_{Adj} III.2 | antépos < **meilleur**_{Adj} II | antépos; **intime**_{Adj} I.1 | postpos, **proche**_{Adj} II | postpos

Qui l’est depuis longtemps
Magn^{temp} **vieil**_{Adj} II | antépos, **de longue date** < **de toujours**

Dont le comportement reflète le caractère fort de la relation
Magn_{comportement} **dévoué** | postpos, **inséparable** II | souvent à propos de \$1 et \$2 avec ~ au pl

Véritable
Ver **véritable** 2, **vrai**_{Adj} II | antépos; **sincère**; **fidèle**_{Adj} 2, **loyal** | postpos; **sûr** II | postpos

Adv₁ **amicalement**

Fait d’être ~
S₀**Pred** **amitié**

[\$1] être ~ de \$2
Oper₁₂ **être**_v II [~ avec N=\$2]

[\$1] commencer à être ~ avec \$2
IncepOper₁₂ **devenir** [~ avec N=\$2]

[\$2] avoir ~
Oper₂ **avoir**_v II [ART ~], **posséder** II [ART ~]

[\$2] cesser d’avoir \$1 comme ~
FinOper₂ **perdre** II [ART ~]

Figure 4 : Extrait de la zone de fonctions lexicales de AM_N I.1 dans le RLF

Nous ne pouvons bien entendu pas entrer dans le détail de l’explicitation de chacune des fonctions lexicales syntagmatiques utilisées ci-dessus, laissant aux formules de vulgarisation le soin de donner au lecteur non-spécialiste une compréhension approximative de la valeur sémantico-syntaxique des liens correspondants.

Outre son utilité pour la description formelle des collocations, le système des fonctions lexicales syntagmatiques joue, pour les linguistes qui l’adoptent, un rôle de grille d’analyse et un rôle prédictif très fort. En effet, on peut d’une certaine façon considérer que ce qui fait d’un syntagme une collocation, encore plus que son caractère semi-phraséologique, c’est le fait même qu’il exprime une fonction lexicale syntagmatique standard. Cette vision des choses peut ainsi pousser le lexicographe à considérer comme collocations, devant potentiellement entrer dans le dictionnaire, des syntagmes que tout un chacun analyserait *a priori* comme syntagmes libres.

Ainsi, on considère habituellement que l’adverbe *très* est en français un intensificateur « joker » pour les adjectifs, ce qui pousse à refuser à un syntagme comme *très amoureux* le statut de collocation. Dans l’approche de la LEC, cependant, on identifie l’expression de l’intensification comme relevant de la fonction lexicale **Magn**, ce qui nous pousse à postuler que l’expression de ce sens auprès d’un adjectif comme *amoureux* relève potentiellement de la phraséologie. En un sens, parce que *très* exprime **Magn** auprès de *amoureux*, il n’est pas forcément évident – par exemple, pour un apprenant du français locuteur d’une autre langue – que le syntagme *très amoureux* est véritablement idiomatique. Dans la mesure où existent des valeurs beaucoup moins prévisibles pour **Magn**(*amoureux*), c’est sans remords qu’une ressource comme le DECFC ou le RLF énuméreront l’ensemble des collocatifs suivants, incluant *très* :

Magn(*amoureux*) = *intensément, profondément, très < comme un fou, éperdument, follement, fou, furieusement, passionnément*

Ne pas inclure *très* dans ce contexte pourrait tout à fait signifier que *très amoureux* n'est pas idiomatique, car, justement, *très* n'est pas compatible avec tous les adjectifs. Ainsi, le syntagme **très mort* est un syntagme à proscrire. La première réaction, dans ce cas-ci, sera bien naturellement de chercher à rationaliser cette impossibilité : on est mort ou pas, on n'est ni un peu mort ni beaucoup. Or c'est faux ! On peut tout à fait dire *raide mort*, ce qui n'est rien de moins qu'une intensification, qui aurait pu en théorie être exprimable par **très mort*.

Mentionnons, pour conclure sur le sujet, que toutes les collocations ne peuvent pas être décrites au moyen des fonctions lexicales standard, car, précisément, certaines collocations ne sont pas standard. Ainsi, *ruisseau capricieux* est de toute évidence une collocation dont *ruisseau* est la base, collocation qui ne peut pas être formalisée dans l'article lexicographique de *RUISSEAU* au moyen d'une formule de fonction lexicale standard. Elle ne peut recevoir qu'une description sous forme de paraphrase, similaire aux formules de vulgarisation présentée ci-dessus ; par exemple : **Qui fait des zigzags**⁸.

3. Phénomène des collocatifs actanciels

Dans les sections précédentes, nous avons replacé la modélisation lexicographique des collocations dans le cadre à la fois plus global et plus précis de la modélisation de la combinatoire restreinte des lexies. Nous sommes maintenant en mesure d'introduire la notion qui est au cœur de la présente discussion : le collocatif actanciel. Nous offrirons tout d'abord une définition et des illustrations de cette notion (section 3.1), pour expliquer ensuite pourquoi une collocation à collocatif actanciel est un syntagme phraséologique de nature hybride (section 3.2).

3.1 Qu'est-ce qu'un collocatif actanciel ?

Les réflexions qui ont précédé nous permettent de définir immédiatement et de façon très simple la notion de collocatif actanciel.

Le collocatif B de la base A dans la collocation AB est un collocatif actanciel s'il exprime un actant sémantique de A dans la collocation en question.

Nous illustrerons cette notion par deux exemples, en commençant par le cas de l'acception du vocable *CRIER* que l'on trouve employée dans une phrase comme *Elle a crié en entendant le coup de feu*, acception que nous numérotions *CRIER I.1*. Il est raisonnable de postuler la structure actancielle suivante pour cette lexie : 'L'individu X crie *I.1* du fait de Y'⁹. Le second actant (Y) de cette lexie s'exprime prototypiquement par un syntagme prépositionnel *de+N* (sans article) : *crier de peur, de joie, de douleur*, etc. On pourrait penser que l'expression de Y auprès de *CRIER I.1* est tout à fait régulière : le sens 'Y' sera exprimé auprès du verbe par *de+L('Y')*, où *L('Y')* représente la ou les lexies signifiant 'Y'. Tout n'est pourtant pas aussi simple. Ainsi, pour lexicaliser Y signifiant 'orgasme', on dira *crier de plaisir* – qui est ambigu, mais veut notamment dire 'crier du fait d'un orgasme' – et aucunement **crier d'orgasme*. Le nom *CRI I.1* correspondant à *CRIER I.1* est, quant à lui, beaucoup plus souple, puisqu'il autorise *cri de plaisir*, mais aussi *cri d'orgasme*. Il est donc clair que *crier de plaisir* et *cri de plaisir/d'orgasme* sont des collocations de, respectivement, *CRIER I.1* et *CRI I.1*. De plus, les collocatifs remplissent ici la position actancielle de second actant des prédicats verbal et nominal : il s'agit donc de collocatifs actanciels.

Pour notre second exemple, prenons le cas de la lexie métaphorique *CEINTURE II* – *Une ceinture de murailles entoure la ville* –, dont la structure actancielle est 'ceinture *II* de X autour de Y'. Le

8 Voir Jousse (2010 : Chapitre 4) pour une réflexion sur le problème de la normalisation des liens de fonctions lexicales non standard.

9 Noter que nous ne traitons pas de *CRIER I.2* (hyponyme de *DIRE*), dont la forme propositionnelle est 'L'individu X crie *I.2* Y à Z'.

premier actant (X) de cette lexie peut s'exprimer par un complément prépositionnel du type *de+N* : *ceinture de murailles*, *ceinture de volcans*, etc. Il peut aussi parfois s'exprimer par des collocatifs adjectivaux épithètes. Ainsi, *ceinture de végétation* a comme paraphrase approximative *ceinture verte*, collocation qui doit clairement être décrite dans l'article lexicographique de CEINTURE II et où le collocatif exprime l'actant X de cette lexie : il s'agit là encore d'un collocatif actanciel.

Le phénomène que nous venons de décrire est loin d'être marginal. Dans chaque langue, il existe un nombre significatif de phrasèmes qui sont des collocations à collocatif actanciel. Il est donc important de savoir bien les identifier et bien les décrire. Cela peut être rendu difficile par le caractère hybride de ces collocations, point que nous allons maintenant discuter.

3.2 Nature hybride des collocations à collocatif actanciel (CCA)

Adoptons la perspective du locuteur – ce qui est un trait caractéristique de l'étude Sens-Texte de la langue, dont relève la LEC – et examinons quel besoin de communication satisfait l'énonciation de *de plaisir* auprès de *crier* ou de *verte* auprès de *ceinture*. Le locuteur qui énonce *crier de plaisir* veut parler d'une action de crier et indiquer que ce qui fait crier X (le « Y de crier ») est un orgasme de X. Il s'agit donc d'exprimer un actant du verbe auprès de ce dernier, en tant que dépendant syntaxique : c'est typiquement ce que met en œuvre le régime d'une lexie. Le même raisonnement pourrait être appliqué à la production de *ceinture verte* par le locuteur.

En d'autres termes, les collocations à collocatif actanciel – désormais, CCA – comme *crier de plaisir* ou *ceinture verte* relèvent en premier lieu du régime syntaxique des lexies qui les gouvernent syntaxiquement et en sont les bases.

Du point de vue lexicographique, décrire de telles collocations relève tout d'abord de la description du régime de leur gouverneur syntaxique. Il n'en demeure pas moins qu'il s'agit tout de même de collocations et que, comme telles, elles relèvent aussi de la description de la combinatoire lexicale restreinte du gouverneur. C'est pour cela que l'on doit considérer que les CCA ont une nature hybride. Dans l'optique de la LEC, un tel phénomène n'est pas surprenant puisque cette approche reconnaît explicitement le caractère non étanche de la séparation entre combinatoire syntaxique restreinte (régime) et combinatoire lexicale restreinte. Il nous semble donc que les CCA, loin d'être des épiphénomènes, sont en réalité une conséquence logique de la perméabilité entre toutes les composantes de la combinatoire restreinte des lexies.

Un autre fait remarquable doit être mentionné à propos de ces phrasèmes. Un syntagme *AB* qui est une CCA est généralement (et peut-être toujours) une collocation du collocatif actanciel lui-même :

- *crier de plaisir* doit être décrit dans l'article lexicographique de CRIER I.1 en tant que CCA de cette lexie ;
- *crier de plaisir* doit aussi être décrit dans l'article lexicographique de la lexie PLAISIR 2 ('plaisir sexuel', cf. *Il n'a jamais connu le plaisir*) comme collocation non-CCA de cette dernière.

Les CCA sont donc, du point de vue lexicographique, ce que l'on pourrait appeler des *collocations réversibles*.

Le caractère à la fois hybride et réversible des CCA peut rendre problématique leur modélisation lexicographique si l'on n'adopte pas à leur propos une stratégie descriptive claire, rigoureuse et systématique. C'est de cela dont il sera maintenant question.

4. Vers une solution lexicographique

Cette dernière section a pour finalité de proposer des stratégies de description lexicographique des

CCA. Nous examinerons successivement deux aspects de la question : CCA et schémas de régime (section 4.1) puis CCA et fonctions lexicales standard (section 4.2).

4.1 CCA et schémas de régime

On doit conclure de la présentation des CCA que nous avons effectuée plus haut que la description de ces collocations doit se fonder, en tout premier lieu, sur la description du régime syntaxique de la lexie base. Il est nécessaire que la structure syntaxique d'une CCA soit « prévue » par le modèle du régime. Ainsi, l'article lexicographique de la lexie CEINTURE II doit indiquer qu'il est possible d'exprimer le premier actant (X) de cette lexie au moyen d'une des deux structures syntaxiques suivantes :

1. une structure de complémentation en *de+N_{pl}* ;
2. une structure modificative adjectivale *Adj*.

Par défaut, le modificateur adjectival *Adj* est un dérivé de *N* : *ceinture de routes ~ ceinture routière*. Cependant, comme nous l'avons mentionné plus haut, il peut aussi s'agir d'un tout autre type d'adjectif : *ceinture de végétation ~ ceinture verte*. Alors que le calcul de la valeur lexicale de *N* est tout à fait libre et peut se faire sur une base entièrement sémantique (*ceinture de biscuits au rhum autour d'une glace à la chantilly, ceinture fers à souder...*, bref, ceinture d'absolument importe quoi), le calcul de la valeur lexicale de *Adj* doit être considéré comme potentiellement non libre, *ceinture II Adj* étant une collocation. On peut essayer d'inclure directement dans le régime syntaxique les lexicalisations possibles de *Adj*, ce qui risque d'alourdir énormément le schéma de régime, ou bien on peut faire suivre le régime de la description des CCA. C'est cette seconde approche qui a été adoptée dans le *Lexique actif du français* (Mel'čuk & Polguère 2007), comme le montre l'extrait de l'article CEINTURE II ci-dessous.

Ceinture où l'on trouve X [= *de+N* (ex. *ceinture d'arbres, de végétation*),
Adj (ex. *ceinture routière*)] et qui se trouve autour de Y [*de N, A_{poss}*]

☞ bande, zone; enceinte; contour, couronne, périphérie, pourtour

Telle que X est 1) de la végétation, 2) un ensemble d'industries, 3) une ou plusieurs voies de circulation importantes, 4) réseau de voies ferrées 1) de végétation; verte | postpos 2) industrielle | postpos 3) routière | postpos 4) ferroviaire | postpos Type particulier de Y agglomération, ville De

Figure 5 : Régime et CCA de CEINTURE II (Mel'čuk & Polguère 2007 : 139)

Apparaît tout d'abord dans la Figure 5 un syntagme qui fournit les différentes positions actancielles contrôlées par la lexie-vedette ainsi que son régime syntaxique. Ce dernier, au lieu d'être présenté dans un tableau à part (comme dans le DECFC – voir Figure 3, section 2.1), est fusionné avec la structure actancielle : chaque mention d'une position actancielle est suivie de l'énumération, entre crochets, des syntagmes pouvant exprimer l'actant en question. La structure syntaxique des CCA de CEINTURE II est donc à ce stade déjà « prévue par le modèle ». Les CCA elles-mêmes sont introduites le plus tôt possible après le régime, juste en dessous de l'énumération des quasi-synonymes (signalés par une petite main pointant à droite). Les CCA sont précédées d'une paraphrase descriptive du type formule de vulgarisation, comme pour tout lien lexical décrit dans le *Lexique actif du français*¹⁰.

La solution adoptée ici rend compte de la nature valencielle des CCA. Pour cela, il y a transgression de l'ordre canonique de structuration en zones d'un article lexicographique construit dans le cadre

¹⁰ Le *Lexique actif du français* est, contrairement au DECFC, un ouvrage destiné aux non-spécialistes. Il n'utilise que les formules de vulgarisation et ne mentionne pas les formules de fonctions lexicales proprement dites.

de la LEC. Cet ordre canonique est : régime, suivi des dérivés sémantiques (fonctions lexicales paradigmatiques), puis des collocations (fonctions lexicales syntagmatiques). Malgré leur nature **formelle** foncièrement syntagmatique, les CCA sont « remontées » dans l'article lexicographique du *Lexique actif du français* auprès du régime, pour mieux refléter leur nature **fonctionnelle**. On pourra cependant reprocher à l'approche adoptée ici de ne pas suffisamment signaler, dans la description même du régime, la nature collocative de l'expression de l'actant X par Adj. La solution descriptive élaborée pour le *Lexique actif du français* laisse donc clairement la place à des améliorations.

Examinons maintenant, pour conclure la discussion, le cas très intéressant des CCA où le collocatif, en plus d'exprimer un actant de la base, possède la valeur sémantico-syntaxique d'une fonction lexicale standard.

4.2 CCA et fonctions lexicales standard

Les CCA *crier de plaisir* et *ceinture verte*, que nous avons examinées plus haut, sont des collocations que l'on pourrait qualifier de non standard, car leur valeur sémantico-syntaxique ne correspond pas à un patron phraséologique récurrent et universel¹¹. Ainsi, si l'on veut décrire *cri de plaisir*, on doit se contenter d'utiliser une paraphrase du syntagme *de plaisir* pour en décrire la valeur, et le système formel des fonctions lexicales standard ne nous est d'aucune utilité. Il faudra donc insérer dans l'article lexicographique de **CRIER I.1** une information du type :

Du fait d'un orgasme Y éprouvé par X : de plaisir

On peut aussi, comme dans le RLF et suivant une pratique introduite dans le cadre de la construction de la base DiCo, formaliser un peu plus les choses :

Y='orgasme de X' : de plaisir

Cette formalisation est toutefois « plate » : elle ne s'insère dans aucun système permettant un calcul logique, contrairement à la formalisation effectuée au moyen des fonctions lexicales standard (Kahane & Polguère 2001). On atteint ici les limites de la description formelle des langues : certains phénomènes linguistiques ne se formalisent tout simplement pas, au sens mathématique.

Ce qui vient d'être dit à propos de nos deux exemples de CCA ne concerne cependant pas toutes les collocations de cette famille. Certaines CCA sont particulièrement intéressantes du fait que, en plus de leurs caractéristiques fonctionnelles remarquables, elles relèvent de fonctions lexicales standard. On peut d'ailleurs faire l'hypothèse que de telles CCA sont présentes en nombre significatif dans toute langue. La valeur la plus typique des collocatifs concernés ici est celle de la fonction lexicale d'intensification, **Magn**, et de son antonyme, **AntiMagn**. Pour expliquer cela, nous nous fondons dans ce qui suit sur une analyse très claire et précise de la combinatoire du verbe **COÛTER I.1** – *Ce sandwich_X a coûté 3€_Y à Paolo_Z* – effectuée dans Elnitsky & Mel'čuk (1988 : 76–78), un des textes introductifs du deuxième volume du DECFC.

Elnitsky & Mel'čuk (1988) observent que le verbe **COÛTER I.1**, tout comme de nombreuses autres lexies dites *paramétriques*, régit des compléments « intensificateurs » de nature métaphorique : *Ça lui a coûté la peau des fesses, un bras, les yeux de la tête*. Pour nous, il s'agit de CCA : les syntagmes nominaux compléments du verbe en sont clairement des collocatifs et ils en expriment son deuxième actant (Y). Une fois que l'on a dit *la peau des fesses*, on ne peut plus exprimer une quelconque autre valeur pour Y. Ou alors, il faut effectuer une coupure syntaxique et prosodique marquant le fait que l'on apporte une précision sur ce qui vient d'être dit : *Ça lui a coûté la peau des fesses, 300€* (= 'Ça lui a coûté la peau des fesses, plus précisément, 300€'). Une CCA de ce type se distingue de celles examinées précédemment par le fait que le collocatif doit être

11 Sur la notion de caractère standard vs non standard d'une fonction lexicale, voir Polguère (2007).

modélisé non seulement comme exprimant un actant de la base (comme dans le cas de *cri de plaisir*), mais aussi comme exprimant un contenu sémantique équivalent à celui de la fonction lexicale **Magn**, ou plus exactement **Magn₂** : intensification portant sur le second actant.

La solution proposée par Elnitsky & Mel'čuk (1988) et implantée dans le DECFC passe par le recours à la fonction lexicale paradigmatique **S₂** en combinaison avec **Magn**, formant ainsi la fonction lexicale paradigmatique complexe **MagnS₂**. Rappelons que **S₂** est la fonction lexicale qui associe à la lexie-vedette la liste de ses dérivés sémantiques dénotant son deuxième actant. Par exemple :

- **S₂**(*boire*) = *boisson*
- **S₂**(*chèque*) = *montant* [de ART ~]

La sémantique de la fonction lexicale **MagnS₂** serait ainsi de dénoter le second actant de lexie-vedette, en tant qu'il exprime l'intensification. La raison pour encoder ainsi la CCA est qu'un encodage immédiatement sous la fonction lexicale **Magn₂** serait formellement invalide : **Magn** dénote un modificateur syntaxique de la lexie-vedette, alors que *la peau des fesses, un bras, etc.*, fonctionnent comme un actant syntaxique (un complément) de celle-ci. Le mérite de cet encodage, outre qu'il permet de respecter des impératifs formels du système des fonctions lexicales, est de faire « remonter » la description de la CCA vers le haut de la zone de fonctions lexicales dans l'article de la lexie-vedette et, donc, de la rapprocher du régime syntaxique, un peu comme dans l'approche adoptée dans le DiCo et le RLF que nous avons introduite un peu plus haut. Notons, finalement, que les observations faites ici valent potentiellement pour n'importe quel actant d'une lexie. On pourrait ainsi considérer tout aussi bien des **MagnS₁**, **MagnS₃**, etc. Ces observations valent aussi pour tout un paradigme de significations de fonctions lexicales syntagmatiques et, à **Magn** et **AntiMagn**, il faudrait ajouter **Ver** (= 'tel qu'il doit être') et **AntiVer**, ainsi que **Bon** (= 'qui est bien') et **AntiBon**.

Nous ne pouvons entrer plus avant dans l'examen de la modélisation proposée dans Elnitsky & Mel'čuk (1988), et nous encourageons le lecteur à se reporter directement au texte en question. La solution avancée, une fois considérée dans le détail, est nous semble-t-il globalement satisfaisante d'un point de vue à la fois théorique et descriptif. Mentionnons toutefois que l'on peut reprocher à la description adoptée dans le DECFC de ne pas insérer le cas particulier des compléments collocatifs des lexies paramétriques dans un cadre plus global, qui est celui des CCA, même si les articles du DECFC contiennent des informations très explicites sur la nature particulière des collocations en question. Ainsi, l'article de **COÛTER I.1** du DECFC contient la description d'un **Magn₂** véritable [*coûter*] *cher* qui a la propriété d'occuper une position actancielle, bien qu'étant un modificateur adverbial : *Ça a coûté cher, 300€* (avec pause indiquant que l'on introduit une précision à propos d'un Y déjà exprimé). Or, justement, un lecteur attentif du DECFC pourra observer que cette propriété d'être un collocatif actanciel est explicitement identifiée : l'adverbe intensificateur apparaît comme option syntaxique, pour l'expression de Y, dans le schéma de régime de la lexie-vedette – cf. Figure 6 ci-dessous¹².

12 La mention du fait qu'il s'agit d'un **Magn₂** ne suffit pas pour indiquer que le collocatif est actanciel. En effet, un **Magn₂** peut tout à fait être non actanciel, comme dans le cas de **Magn₂**(*CHÈQUE* [signé par X pour un montant Y à Z à partir du compte W]) = *gros*. L'expression de l'intensification au moyen de l'adjectif épithète *gros* ne sature aucunement la position actancielle Y : *Il lui a signé un gros chèque de 300€*.

1 = X	2 = Y
1. N 2. V _{inf}	1. Num N 2. N 3. Adv obligatoire

- 1) C_{2,2} : N = *prix I, somme, MagnS₂, AntiMagnS₂*
- 2) C_{2,3} : Adv = *Magn₂, AntiMagn₂, combien*
- C₁ + C₂ : *Ce livre coûte 10 francs <un prix exorbitant, cher>; Ce livre ne coûte qu'une bouchée de pain; Aller à Paris coûte une fortune*

Figure 6 : Régime de COÛTER **I.1** dans Mel'čuk et coll. (1988 : 160) = Volume II du DECF

En conclusion, nous résumerons en quatre points les solutions proposées pour la modélisation lexicographique des CCA.

- 1) La structure des CCA doit être prévue dans le schéma de régime syntaxique de la lexie qui en est la base, avec mention éventuelle du fait qu'une construction donnée, présente dans le régime, est réservée exclusivement à des collocatifs actanciels.
- 2) La description des CCA qui ne relèvent pas de fonctions lexicales standard apparaît en tant qu'ajout au régime, au même titre que les contraintes de réalisation situées sous le schéma de régime dans les entrées du DECF.
- 3) La description des autres CCA sera faite au moyen de formules de fonctions lexicales standard du type **MagnS₂**, **Magn₂**, etc., accompagnées de formules de vulgarisation « grand public ».
- 4) Bien entendu, puisque les CCA sont des collocations réversibles (cf. section 3.2), elles pourront s'il le faut être aussi décrites dans l'article de la lexie fonctionnant comme collocatif au sein de la CCA. Dans cet encodage inversé, la base sera alors le dépendant syntaxique et le collocatif le gouverneur. Par exemple, **CRIER I.1** sera aussi décrit comme collocatif de la lexie **PLAISIR 2** (sexuel) dans l'article de cette dernière :

[X] **produire un son exprimant son** ~ : *crierI.1* [de ~]

Nous formulons bien évidemment nos propositions en nous appuyant sur les outils formels disponibles dans la LEC (tout particulièrement, les schémas de régime et le système des fonctions lexicales). Il faut cependant attacher de l'importance avant tout à l'éclairage que de telles analyses apportent sur le phénomène des CCA. Les descriptions proposées peuvent tout à fait être traduites dans d'autres systèmes d'encodage lexicographique. Ce qui compte, c'est plus la conceptualisation des CCA permise par la modélisation que la forme de cette modélisation elle-même. Cela est particulièrement important, notamment, dans le contexte d'une exploitation pédagogique des pistes descriptives introduites ici.

Bibliographie

Baker, Collin F., Fillmore, Charles J., Cronin, Beau. 2003. The Structure of the FrameNet Database. *International Journal of Lexicography* 16(3), 281–296.

- Elnitsky, Léo, Mel'čuk, Igor. 1988. le champ sémantique 'grandeurs' (= 'paramètres') : description lexicographique de quelques cas problématiques de la cooccurrence lexicale (actants à valeur numérique et modificateurs de degré). Dans Mel'čuk et coll. (dir.) (1988), 73–80.
- Fillmore, Charles S. 1982. Frame Semantics. Dans Linguistic Society of Korea (dir.) : *Linguistics in the Morning Calm. Selected papers from SICOL-198*, Séoul, Hanshin Publishing Co., 111–137.
- GDEL. 1982–1985. *Grand dictionnaire encyclopédique Larousse : GDEL*. 10 volumes, Paris, Larousse.
- Gross, Gaston. 2008. Les classes d'objets. *Lalies* 28, 111–165.
- Gross, Maurice. 1975. *Méthodes en syntaxe. Le régime des constructions complétives*, Paris, Herman.
- Hausmann, Franz Josef. 1989. Le dictionnaire de collocations. Dans Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand & Ladislav Zgusta (dir.) : *Wörterbücher : ein internationales Handbuch zur Lexicographie / Dictionaries: An International Encyclopaedia of Lexicography / Dictionnaires : Encyclopédie internationale de lexicographie*, Handbücher zur Sprach- und Kommunikationswissenschaft, Band 5.1, Berlin/New York NY, Walter de Gruyter, 1010–1019.
- Herbst, Tomas. 1988. A Valency Model for Nouns in English. *Journal of Linguistics* 24(2), 265–301.
- Herbst, Tomas., Heath, David, Roe, Ian F., Götz, Dieter. 2004. *A Valency Dictionary of English : A Corpus-Based Analysis of the Complementation Patterns of English Verbs, Nouns, and Adjectives*, Topics in English Linguistics 40, Berlin/New York, Walter de Gruyter.
- Iordanskaja, Lidija, Mel'čuk, Igor. 2009. Establishing an Inventory of Surface-Syntactic Relations : Valence-Controlled Surface-Syntactic Dependents of the Verb in French. Dans Alain Polguère & Igor Mel'čuk (dir.) : *Dependency in Linguistic Description*, Studies in Language Companion Series 111, Amsterdam/Philadelphia, 151–234.
- Jousse, Anne-Laure. 2010. *Modèle de structuration des relations lexicales fondé sur le formalisme des fonctions lexicales*. Thèse de doctorat, Département de linguistique et de traduction – Université de Montréal & UFR de linguistique – Université Paris Diderot (Paris 7).
- Kahane, Sylvain, Polguère, Alain. 2001. Formal Foundation of Lexical Functions. *Proceedings of COLLOCATION : Computational Extraction, Analysis and Exploitation*, 39th Annual Meeting and 10th Conference of the European Chapter of the Association for Computational Linguistics, Toulouse, 8–15.
- Lux-Pogodalla, Veronika, Polguère, Alain. 2011. Construction of a French Lexical Network : Methodological Issues. *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop*, Ljubljana, 54–61.
- Mel'čuk, Igor. 1995. Phrasemes in Language and Phraseology in Linguistics. Dans Martin Everaert, Erik-Jan van der Linden, André Schenk & Rob Schreuder (dir.) : *Idioms: Structural and Psychological Perspectives*, Hillsdale N.J. 7/Hove UK, Laurence Erlbaum Associates, 167–232.
- Mel'čuk, Igor. 1996. Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon. Dans Leo Wanner (dir.) : *Lexical Functions in Lexicography and Natural Language Processing*, Language Companion Series 31, Amsterdam/Philadelphia, John Benjamins, 37–102.
- Mel'čuk, Igor. 2004a. Actants in semantics and syntax I : actants in semantics. *Linguistics* 42 : 1, 1–66.
- Mel'čuk, Igor. 2004b. Actants in semantics and syntax II : actants in syntax. *Linguistics* 42 : 2, 247–291.

- Mel'čuk, Igor et coll. 1984, 1988, 1992, 1999. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I-IV*, Montréal, Les Presses de l'Université de Montréal.
- Mel'čuk, Igor, Clas, André, Polguère, Alain. 1995. *Introduction à la lexicologie explicative et combinatoire*, Paris/Louvain-la-Neuve, Duculot.
- Mel'čuk, Igor, Polguère, Alain. 2006. Dérivations sémantiques et collocations dans le DiCo/LAF. *Langue française* 150, 66–83.
- Mel'čuk, Igor, Polguère, Alain. 2007. *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*, Champs linguistiques, Bruxelles, De Boeck & Larcier.
- Milićević, Jasmina. 2009. Schéma de régime : le pont entre le lexique et la grammaire. *Langages* 176, 94–116.
- Moon, Rosamund. 2008. Dictionaries and collocation. Dans Sylviane Granger & Fanny Meunier (dir.) : *Phraseology. An interdisciplinary perspective*, Amsterdam/Philadelphia, John Benjamins, 313–336.
- Polguère, Alain. 2007. Lexical function standardness. Dans Leo Wanner (dir.) : *Selected Lexical and Grammatical Issues in the Meaning-Text Theory. In Honour of Igor Mel'čuk*, Language Companion Series 84, Amsterdam/Philadelphia, John Benjamins, 43–95.
- Sinclair, John M. (dir.) 1987. *Collins COBUILD English language dictionary*, London, Collins.
- Van den Eynde, Karel. 1998. From Verbal to Nominal Valency : Some Methodological Reflections. Dans Karen Van Durme and Lene Schøsler (dir.) : *Studies in Valency IV. Valency and Verb Typology*, Odense, Odense University Press, 147–167.
- Van den Eynde, Karel, Miertens, Piet. 2003. La valence : l'approche pronominale et son application au lexique verbal. *Journal of French Language Studies* 13, 63–104.
- Van den Eynde, Karel, Miertens, Piet. 2006. *Le dictionnaire de valence DICOVALENCE : manuel d'utilisation*. Université de Leuven.