



**HAL**  
open science

# Quelques questions sur le rapport entre la linguistique fondée sur l'usage et l'organisation cognitive de la compétence

Dominique Legallois

## ► To cite this version:

Dominique Legallois. Quelques questions sur le rapport entre la linguistique fondée sur l'usage et l'organisation cognitive de la compétence. Usages, vol. 6,, Iste éditions, inPress, Les concepts fondateurs de la philosophie du langage. hal-02055187

**HAL Id: hal-02055187**

**<https://hal.science/hal-02055187>**

Submitted on 3 Mar 2019

**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.

## Chapitre 11

### Quelques questions sur le rapport entre la *linguistique fondée sur l'usage* et l'organisation cognitive de la compétence

On peut mesurer l'importance actuelle de la notion de *linguistique fondée sur l'usage* à la fréquence du terme employé dans les titres et sous-titres d'ouvrages récemment publiés ; en voici quelques exemples, parmi une multitude possible :

*Usage-based approaches to Dutch : lexicon, grammar, discourse* [VER 03] ; *English Ditransitive Verbs : Aspects of Theory, Description and a Usage-Based Model* [MUK 05] ; *Constructing a Language: A Usage-Based Theory of Language Acquisition* [TOM 05] ; *Non-finite complementation : A usage-based study of infinitive and -ing clauses in English* [EGA 08] ; *Germanic Future Constructions : A Usage-Based Approach to Language Change* [HIL 08] ; *Rethinking idiomaticity : a usage-based approach* [WUL 10] ; *Entrenchment in Usage-Based Theories : What Corpus Data Do and Do Not Reveal about the Mind* [BLU 12] ; *Usage-based Approaches to Japanese Grammar : Towards the understanding of human language* [KAB 14] ; *Understanding Relative Clauses : A Usage-Based View on the Processing of Complex Constructions* [WIE 14] ; *Preposition Placement in English : A Usage-based Approach* [HOF 14] ; *The Usage-based Study of Language Learning and Multilingualism* [ORT 16] ; etc.

Au regard de cet échantillon, la notion d'usage caractérise des études sur des phénomènes linguistiques généraux : il s'agit a minima de signaler une méthodologie. Cette méthodologie requiert généralement une investigation sur corpus ou sur d'autres types de données empiriques (bases de données, expériences

psycholinguistiques). Cependant, les études fondées sur l'usage ne se confondent pas entièrement avec ce type de méthodologie : premièrement, le nombre d'articles ou de livres comprenant le terme *corpus-based* est encore bien plus important dans la littérature que *usage-based*, ce qui montrerait que les deux termes ne sont pas tout à fait équivalents et que le second posséderait un champ d'application plus restreint ; deuxièmement, *usage-based* modifie également des termes comme *approach* (qui est sans doute ambigu quant à son contenu : approche méthodologique ou théorique ?) et surtout *view*, *model* et *theory* : l'usage renvoie visiblement à une conception de la langue ou à un ensemble de théories linguistiques et ne se limite donc pas à la seule perspective méthodologique.

L'objectif de ce travail, complémentaire de l'article de D. Legallois et J. François [LEG 11]<sup>1</sup>, est de mettre en perspective cette conception de la langue en présentant ses principaux aspects, ses ramifications, mais aussi les difficultés qu'elle soulève. Pour ce faire, nous consacrons la première partie à la présentation des principaux concepts gravitant autour de la linguistique fondée sur l'usage (désormais LU) : on doit considérer que l'expression de LU est en quelque sorte une étiquette dénommant un groupement de notions. C'est seulement à travers ces notions, proposées principalement par la Linguistique Cognitive de R. Langacker (par exemple, le rapport schématicité-spécificité, la fréquence, la routinisation, l'ancrage, etc.) qu'il est possible, selon nous, de définir véritablement LU. La deuxième partie aborde la conception *usage-based* en acquisition, principalement dans le modèle de M. Tomasello. La troisième partie porte sur les niveaux de catégorisation des unités linguistiques et présente un débat très actuel en LU : la complémentarité et l'opposition entre les phénomènes d'unités prototypiques de l'usage, et les phénomènes d'exemplaires. On y discute également des problèmes à concevoir une compétence statistique des locuteurs. La dernière partie porte sur le rapport entre LU et grammaire du discours (terme emprunté à J.P. Sueur) ; on y montre la complexité des connaissances liées à l'usage : comment transposer en compétence cognitive un savoir d'ordre différent (linguistique, textuel, générique et social) ?

Le parcours proposé est nécessairement fragmentaire. Mais la limite dont nous voulons avertir le lecteur ici n'est pas tant dans l'aspect partiel et partial de l'étude que dans la définition de son objet : celui-ci est consacré à une théorie linguistique, donc la linguistique fondée sur l'usage, et non pas à la notion d'usage en linguistique, qui constituerait un travail quelque peu différent. En effet, comme déjà évoqué plus haut, la LU est une théorie ou une approche de la langue, qui possède son corps d'éléments notionnels ; la notion même d'usage, elle, est à situer dans le

---

<sup>1</sup> Nous renvoyons également à l'article très documenté de [CHE 12] qui, de plus, propose un rapprochement intéressant entre la notion de compétence chez Hymes et les travaux de la linguistique fondée sur l'usage.

débat entre *analogie* (conformation au système) et *anomalie* (l'écart constaté en discours), ou bien encore entre la *norme-règle* et la *norme-habitude*. Ce travail ne traite donc pas de ce dernier objet.

### 11.1. La LU : approche théorique

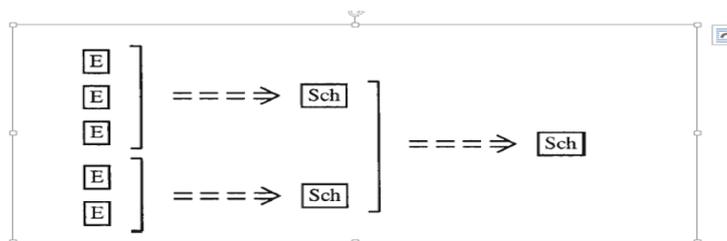
La paternité de la LU est signée : le linguiste américain R. Langacker crée le terme de *modèle fondé sur l'usage* (*usage-based model*) dans le premier tome des *Foundations of Cognitive Grammar* en 1987 [LAN 87]. L'histoire des raisons scientifiques qui ont amené l'initiateur de la Cognitive Grammar à adopter un modèle profondément empirique, non dérivationnel et maximaliste, en réaction contre la Grammaire Générative, est excellemment retracée par J. M. Fortis [FOR 11] ; nous renvoyons donc à cet article pour l'exposé des motivations théoriques, mais il convient cependant de conserver à l'esprit la dimension antagoniste *Linguistique Cognitive vs Linguistique Générative* qui explique les notions associées à la LU. Elle constitue le point de départ de la réflexion de Langacker dans un article [LAN 00] publié dans l'ouvrage collectif *Usage-based models of language* édité par les linguistes américains M. Barlow et S. Kemmer<sup>2</sup>. Ainsi, la Grammaire Générative, plus encore avec la version minimaliste, est une approche fondée sur l'économie des connaissances qu'un locuteur doit posséder pour parler sa langue. On connaît l'axiome : un minimum d'opérations produit une infinitude d'énoncés. Dans ce cadre, ce que l'on appelle traditionnellement les règles et les unités (ou si on veut, la syntaxe et le lexique) constituent des dimensions différentes et exclusives (conception que Langacker qualifie de *rule/list fallacy*). Cette approche formelle, qui se distingue assez peu sur ce point de la grammaire traditionnelle, peut être appréhendée par la métaphore du dictionnaire et de l'ouvrage de grammaire<sup>3</sup> : deux ouvrages de référence distincts, qui possèdent chacun leur domaine, leur organisation et s'associent seulement en production d'énoncés par une *interface*. Ainsi, l'approche formelle que constitue la Grammaire Générative fait une distinction forte entre compétence linguistique des locuteurs et performance. La Grammaire Cognitive défend l'idée que les unités et les structures linguistiques émergent à partir des emplois en discours. Elle est maximaliste, non seulement parce qu'elle ne considère pas la distinction compétence / performance comme pertinente (la compétence est en quelque sorte un savoir performer), mais aussi parce qu'elle admet la redondance des informations linguistiques à plusieurs niveaux d'appréhension de la grammaire ; ainsi, le lexique, les expressions et les « règles »

---

<sup>2</sup>L'ouvrage constitue encore la référence principale à la LU.

<sup>3</sup>Nous empruntons cette métaphore à J. Taylor [TAY 15].

grammaticales forment ensemble un continuum dans le *répertoire*<sup>4</sup> des locuteurs. Ce répertoire, dépositaire de la compétence linguistique, est organisé en réseau : les unités lexicales sont reliées aux diverses constructions auxquelles elles participent, et ces constructions sont appréhendables à des niveaux de spécificité / schématicité différents : par exemple, des séquences figées (le pôle « + spécifique ») aux séquences les plus schématiques, c'est-à-dire sans saturation lexicale (par ex. :  $X_s V Z_o$  à  $Y_{oi}$  – la construction à double objet direct et indirect en français serait enregistrée comme un tout, à l'image d'un signe lexical). Dans cette architecture grammaticale, l'importance de l'opération cognitive de schématisation est à souligner. Les locuteurs abstraient des formes à partir d'événements d'usage (*usage events*), c'est-à-dire des énonciations nécessairement hétérogènes. Ces formes peuvent faire l'objet d'abstractions supplémentaires, comme l'indique la figure ci-dessous, où E représente les événements d'usage et Sch les schémas abstraits :



-Fig. 1 Schématisation d'après [LAN 87]

Le même processus, appliqué à un exemple linguistique (l'acquisition par les enfants des formes interrogatives avec modaux en anglais) :

<sup>4</sup> C'est ainsi que nous traduisons *repository*. La métaphore du répertoire s'oppose donc à celle du dictionnaire / grammaire, en ce sens où le répertoire comprend en même temps la dimension lexicale et la dimension syntaxique.

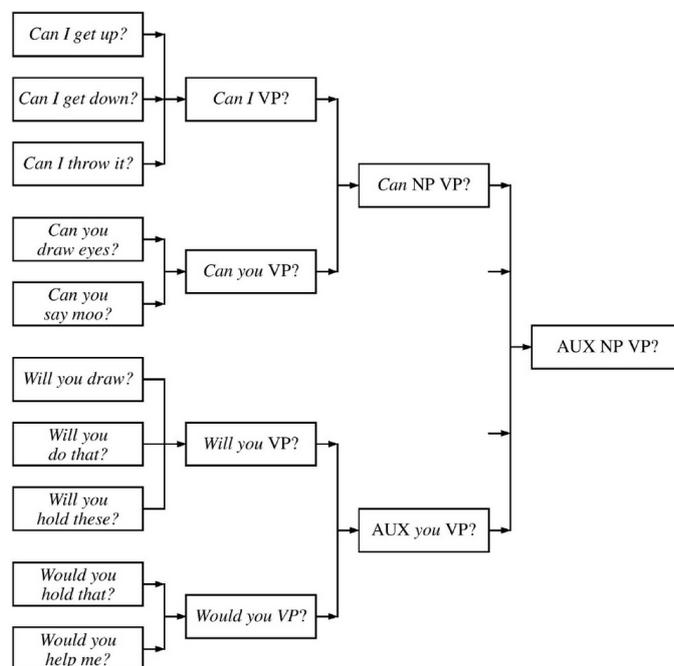


Fig. 2 Exemple de schématisation progressive<sup>5</sup>

Le rôle de la fréquence des unités est déterminant pour leur schématisation ; plus une unité est « rencontrée », plus son appartenance au système linguistique d'un individu est grande. La place des unités est ainsi inégalitaire dans le répertoire : non seulement il est constitué d'unités de natures différentes (des unités discrètes aux schémas), non seulement les unités entretiennent entre elles des relations complexes, mais encore faut-il considérer que les unités se différencient par des poids et donc des importances différents – phénomène que les approches formelles ignorent. C'est bien sûr l'usage qui est responsable de ce déséquilibre entre unités<sup>6</sup>. Le répertoire étant dynamique, puisque l'expérience discursive des locuteurs est permanente, la place de chaque unité dans le réseau se distend, ou au contraire se renforce, selon les événements d'usage. Le phénomène d'*entrenchment* – terme que nous traduirons par *ancrage (cognitif)* – correspond au poids de la place de l'unité dans la compétence

<sup>5</sup> Représentation empruntée au site : <https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00852/full>

<sup>6</sup> Cette conception n'est pas sans rappeler les propos de P. Guiraud [GUI 54] au sujet du lexique : la fréquence, où le sentiment de fréquence, est un trait inhérent à toute unité lexicale.

du locuteur<sup>7</sup> ; plus une unité est ancrée, plus les automatismes, les routines, et les formations d'habitude sont considérables et se renforcent, à l'image des circuits neuronaux des modèles connexionnistes. L'unité et les structures qui lui sont associées deviennent, par ancrage, plus facilement « manipulables », sans qu'il y ait besoin d'une attention consciente de la part du locuteur (et des interlocuteurs) envers la dimension compositionnelle de l'emploi de l'unité.

Par ailleurs, R. Langacker [LAN 00] identifie le rôle joué par certaines capacités psychologiques en LU. Ainsi, la représentation de la schématisation présuppose la mobilisation de capacités cognitives générales ; la LU plaide donc contre la modularité de l'esprit, et le cloisonnement des facultés. En effet, la schématisation est une opération particulière de catégorisation qui, par abstraction, filtre ce qui n'est pas récurrent dans l'expérience, et fait émerger ce qui est commun<sup>8</sup>. Il s'agit également de pouvoir comparer plusieurs expériences et de détecter des différences entre structures, certaines fonctionnant comme standards, les autres comme cibles. La catégorisation et la schématisation reposent donc en partie sur la faculté de comparer. Abstraction et comparaison sont deux opérations cognitives qui ne sont pas exclusivement dédiées au langage ; le « matériau » linguistique appréhendé en usage constitue ainsi une donnée sur lequel des capacités générales s'appliquent. Minimisant ou niant l'importance de l'usage, les approches formelles ne peuvent recourir qu'à des opérations spécifiques au traitement linguistique pour expliquer le fonctionnement grammatical d'une langue.

## 11.2. La LU en acquisition

Modèle cognitif et non essentiellement linguistique, La LU inspire mais aussi a été inspirée par les études sur l'acquisition du langage. On peut esquisser ainsi les grandes lignes des résultats issus des travaux du psycholinguiste M. Tomasello. Selon la Grammaire Générative, qui continue à avoir une grande influence dans le milieu de la psycholinguistique, lorsqu'un enfant maîtrise une structure, par exemple le passif, il est capable de l'appliquer de manière générale et automatique sur toute unité. En LU, au contraire, l'acquisition d'une structure grammaticale est hautement dépendante des données linguistiques auxquelles l'enfant a été confronté. C'est à partir d'un niveau significatif d'exposition que les généralisations apparaissent. Le développement est donc non linéaire. Ainsi, il ressort des études de M. Tomasello [TOM 03] que les enfants commencent par produire des énoncés que l'on peut qualifier de polylexicaux entre l'âge de 18 et 24 mois. A ce stade, on ne peut pas

---

<sup>7</sup> Voir la présentation de Ph. Gréa dans [LEG 06]. Le rapport entre entrenchment et analyse de corpus fait l'objet de l'ouvrage de [BLU 12].

<sup>8</sup> Cf. [LEG à paraître] sur l'importance de la schématisation en linguistique.

parler de syntaxe ; les énoncés en question sont des productions, après les hologrammes, de deux / trois mots, des sortes de collocations telles que *veux biberon*, *veux maman*, *encore eau*, *encore biberon*, etc. Certaines de ces collocations vont former des « patrons », où ce que Braine (1963) a appelé des schémas pivots<sup>9</sup> (*pivot schemas*) et que Tomasello nomme *utterance schema* : des structures avec un mot pivot telles que *veux N*, *encore N* accueillant de nouveaux items lexicaux. Ensuite, intervient un stade où la syntaxe apparaît, mais de manière très particulière : la compétence s'enrichit de constructions fondées sur un seul item (*item-based constructions*), autrement dit, chaque verbe (pour prendre le cas de cette catégorie) possède sa grammaire, indépendamment des autres verbes. Il y aurait une sorte d'idiomaticité radicale des unités. Il s'agit là de l'hypothèse de « l'îlot verbal »<sup>10</sup> ; la grammaire des enfants serait en partie « insulaire ». A un niveau plus avancé du développement linguistique, à 3 ans, l'enfant devient capable de procéder à des généralisations par schématisation et par analogie. Il construit alors peu à peu un système dans lequel les unités sont connectées les unes aux autres, et où sont abstraits des schémas grammaticaux. Par exemple, dans les deux constructions '*X is Y-ing the Z*' and '*A is B-ing the C*', les relations entre les éléments sont les mêmes dans les deux cas : X est analogue à A, Y est analogue à B, Z à C. L'enfant peut ainsi identifier des constructions schématiques et développer ses compétences grammaticales.

Le développement, très rapidement et partiellement présenté ici, ne pourrait avoir lieu sans des capacités mobilisées dans d'autres types d'expériences ; parmi celles-ci<sup>11</sup>, figure la capacité vers 14 mois à concevoir mentalement une procédure à multiples phases pour satisfaire un objectif ; cette faculté serait un prérequis nécessaire pour opérer ensuite des combinaisons de mots. Une autre faculté indispensable, notamment pour former des schémas pivots, est l'instanciation de schémas sensorimoteurs, comme la possibilité de procéder aux mêmes actions avec des objets différents ; on retrouve au niveau linguistique cette capacité lorsque l'enfant applique un même patron à des unités linguistiques différentes.

### 11.3. LU et niveaux de catégorisation

La redondance, qui n'a jamais bonne presse, est revendiquée en LU : une même structure peut être catégorisée à un niveau de base et à un niveau superordonné. Par exemple,

<sup>9</sup> Et qui constituent donc une *pivot grammar*.

<sup>10</sup> Cf. [MIN 03] pour une critique de cette hypothèse.

<sup>11</sup> Nous ne parlons pas ici la capacité de l'attention conjointe, fondamentale pour Tomasello pour comprendre les aspects spécifiquement humains de la cognition.

*L'objectif est de + Inf.* (l'objectif est de faire baisser le chômage)

est une unité semi-spécifique (elle possède une dimension lexicalisée et une dimension ouverte, du même ordre, d'ailleurs, que les schémas pivots) qui émerge de l'usage et constitue ainsi une routine discursive ; elle est également une instance d'un schéma superordonné :

*Det N est de + Inf.*

qui est la catégorie supérieure d'autres instances comme *l'ambition est de Inf.* ; *la possibilité est de Inf.*, *la solution est de + Inf.*, etc.

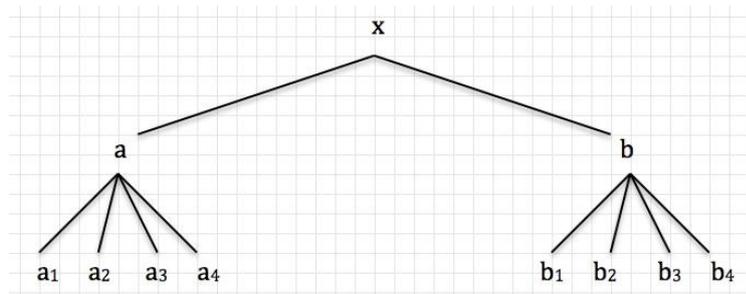
Ce rapport hiérarchique de niveaux de catégorisation ou de schématisation différents est une question fort complexe : cognitivement, le rapport est certainement pertinent, à savoir que les locuteurs adultes peuvent théoriquement procéder à ces multiples catégorisations, puisqu'elles sont banales et normales pour d'autres objets non linguistiques (un teckel est un chien qui est un animal) ; du point de vue des données en corpus, on peut mettre en évidence les liens entre un schéma comme *l'objectif est de + Inf.* et de nombreux autres schémas moins fréquents, qui laissent à penser qu'il existe une configuration plus abstraite. Il reste néanmoins que le rapport entre les niveaux fait l'objet de discussion constante en LU : où, exactement, se « fait » la grammaire d'un locuteur ? A. Goldberg [GOL 06] considère l'importance du niveau spécifique pour la connaissance et la maîtrise des structures (cf. également [BYB 10] ; mais la production d'énoncés nouveaux nécessite également la connaissance d'unités d'ordre schématique : par exemple, dans cet énoncé attesté :

« On adore cuisiner. Ça fait 20 ans qu'on avait envie d'ouvrir un restaurant ». Carole Bussy conjugue cette fièvre au pluriel pour la partager avec Arnaud, son mari. Le seul *antipyrétique* du couple était de trouver un endroit » (*L'Est Républicain*).

*Antipyrétique*, de toute façon peu fréquent, est un nom non prédestiné dans la construction ; son emploi présuppose, premièrement, que le schéma *Det N est de + Inf.* fasse partie de la compétence grammaticale du locuteur, deuxièmement, que le lecteur soit capable d'identifier une analogie, certainement avec le patron « la solution est de + Inf. ». Il est assez raisonnable de penser que des différences interindividuelles de compétence (et donc de performance) linguistique résident dans la maîtrise et la manipulation d'unités de haut niveau.

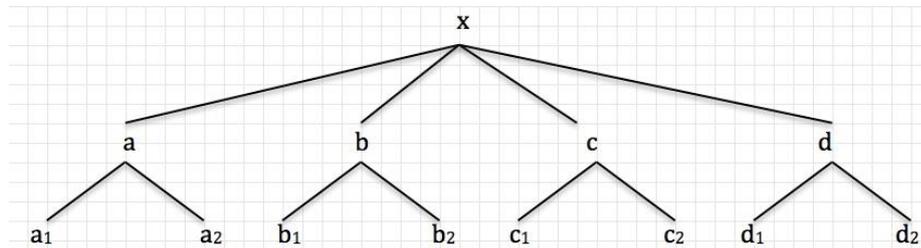
La fréquence étant une donnée essentielle, quelques principes se dégagent permettant de prédire le degré de schématisation de certaines structures. Dans le cas où une structure (X) possède un nombre d'instances limité, si ces instances (a, b)

sont cependant fréquentes ( $a_1..a_4$ ,  $b_1..b_4$ ), X est alors peu enclin à généralisation ; soit la représentation :



-Fig. 3 Conditions pour une non généralisation de X

Si une structure (X) possède de nombreuses instances (a, b, c, d), et si ces instances sont peu fréquentes, alors X est plus facilement généralisable, bien que les instances (en raison de leur nombre) soient chacune difficilement mémorisables. Soit la représentation :



-Fig. 4 Conditions pour une généralisation de X

Pour rendre plus clairs les problèmes que pose le rapport entre niveaux inférieurs et niveaux supérieurs en LU, nous prendrons un exemple. Insistons sur le fait que cet exemple n'a de valeur ici qu'illustrative, et permet de mieux raisonner à partir d'un cas concret. Nous ne prétendons nullement apporter des réponses aux questions soulevées.

Il existe en français (comme en anglais) trois constructions que l'on peut qualifier de « spécificionnelles » : elles constituent des dispositifs de spécification du contenu nominal d'un nom dit « sous-spécifié » ( $N_{ss}$ ) par un contenu propositionnel, infinitival ou nominal :

*$N_{ss}$  être de Inf.* (l'objectif est de faire diminuer le chômage)

$N_{ss}$  **être que P.** (*la conséquence est que le chômage augmente*)

$N_{ss}$  **être N** (*le problème est le chômage*)

Les noms entrant dans ces structures expriment six types de modalités présentées ci-dessous. Nous donnons quelques exemples de noms :

L'aléthique, exprimant la vérité présentée dans ce qu'elle a d'objectif : fait, évidence, logique, vérité, réalité, preuve, résultat, effet, conséquence, raison, motif, etc.

L'épistémique, exprimant la vérité d'un locuteur : sentiment, certitude, crainte, décision, conception, conviction, philosophie, politique, question, réponse, etc.

L'appréciatif, exprimant un jugement subjectif, selon ce qui est désirable : comble, essentiel, important, inconvenient, intérêt, mieux, pire, erreur, scandale, problème, difficulté, risque, solution, etc.

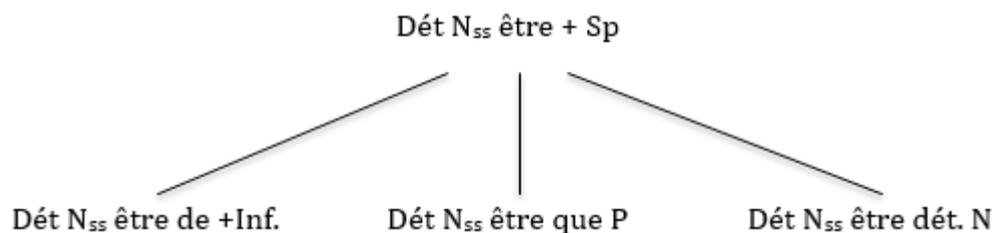
L'axiologique, exprimant un jugement de valeur de nature morale, idéologique ou légale : courage, honneur, mérite, sagesse, vertu, habileté, faute, etc.

Le boulique, exprimant le désir, la volonté, le souhait : ambition, but, choix, intention, objectif, objet, désir, souhait, volonté, priorité, projet, idée, etc.

Le déontique, exprimant l'obligatoire, l'interdit, le permis ou le facultatif, selon les institutions : devoir, responsabilité, mission, rôle, principe, etc.

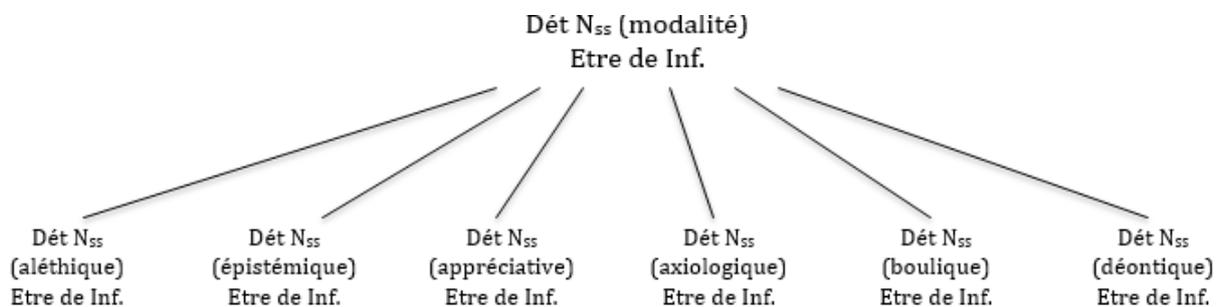
Du point de vue de la LU, les questions suivantes se posent :

A- Existe-t-il une construction spécificationnelle générique (schématique) ayant pour réalisations possibles la construction complétive, la construction infinitive, la construction nominale ? On serait ici dans une architecture avec schématicité maximale. A cela s'ajouteraient dans l'organisation en réseau, les relations avec d'autres constructions, comme la construction pseudo-clivée, indiscutablement apparentée à la construction spécificationnelle et à ses réalisations (par ex. : *ce qui pose problème / le problème*). Soit la représentation :



- Fig. 5 Schématisation maximale de la construction spécificionnelle

B- La distinction entre les modalités, si elle est fondée d'un point de vue sémantique, l'est-elle du point de vue cognitif ? Autrement dit, cette distinction fait-elle partie du savoir lexico-grammatical des locuteurs ? W. Croft [CRO 03, p. 57-58] pose ainsi qu'il existe des constructions qui ne seraient pas entièrement schématiques, mais seraient associées à des classes lexicales. Pour reprendre notre exemple, il pourrait exister dans la compétence linguistique des locuteurs, une construction spécificionnelle avec infinitive associée avec les noms sous-spécifiés de la catégorie appréciative, une autre serait associée avec l'aléthique, une autre encore avec l'appréciatif, etc., et ainsi pour les trois variantes. Soit la représentation :



- Fig. 6 Association entre constructions et classes modales

C- S'il existe des constructions associées à des classes de noms, comment ces classes sont-elles organisées ? Là encore, il existe plusieurs façons d'envisager la catégorisation :

a) soit on considère que la classe s'organise autour d'un prototype, qui pourrait être le nom le plus fréquemment employé, ou plutôt, le plus attiré par la construction. Dans un corpus de débats à l'Assemblée Nationale (composé de 280000 interventions et de 3 874302 mots), nous mesurons la part que les noms de

modalité appréciative et ceux de modalité boulique consacrent à la construction  $N_{ss}$  *être de Inf* (par rapport à leur emploi dans les autres constructions). Sans donner ici de chiffres<sup>12</sup>, on peut proposer une représentation en nuages de mots, non pas pour l'esthétique de la chose, mais parce que ce type de figuration permet de saisir directement les items les plus représentatifs de chaque catégorie (la disposition et la proximité des noms sur ces figures sont aléatoires, ce qui importe est la taille des items). Pour la construction avec les noms de modalité appréciative :



- Fig. 7 Nuage des noms appréciatifs attirés par la construction

Pour les emplois bouliques :



- Fig. 8 Nuage des noms bouliques attirés par la construction

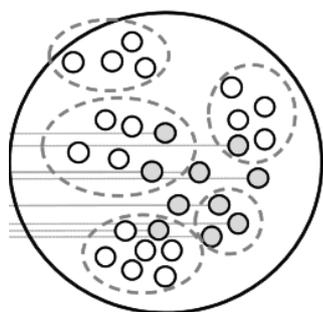
La catégorisation d'un nouvel élément lexical requiert sa comparaison avec le prototype. Cependant, cette conception de la catégorisation autour d'un prototype est concurrencée par une autre vision qui consiste à considérer que les catégories comprennent non pas des unités abstraites de l'expérience, mais directement les

<sup>12</sup> Voir pour cela [ADL à paraître].

exemplaires rencontrés dans l'expérience. Cette thèse<sup>13</sup>, formulée à partir des travaux des psychologues D. Hintzman [HIN 86] et R. Nosofsky [NOS 92] notamment, est principalement défendue en phonologie (par ex. [PIE 01] : alors que les prototypes sont des « types », la théorie des exemplaires propose de concevoir que les locuteurs acquièrent les catégories phonologiques en se remémorant des « tokens », c'est-à-dire les événements d'usage au cours desquels une variété de prononciation d'un même phonème a eu lieu. Aussi, ce n'est pas seulement l'information phonétique qui est emmagasinée, c'est aussi l'information sémantique et plus largement contextuelle. Dans ce cadre (dit *exemplar-based model*), une catégorie est donc définie par une collection d'exemplaires concrets ou de tokens mémorisés. Voici la définition qu'en donne J. Pierrehumbert :

« In an exemplar model, each category is represented in memory by a large cloud of remembered tokens of that category. These memories are organized in a cognitive map, so that memories of highly similar instances are close to each other and memories of dissimilar instances are far apart... When a new token is encountered, it is classified... according to its similarity to the exemplars » [PIE 01, p. 140–141]<sup>14</sup>.

Tout exemplaire peut apparaître simultanément dans plusieurs catégories. Voici une représentation possible de l'organisation des catégories par exemplaires ; les billes grisées représentent de nouvelles occurrences qui sont catégorisées dans des sous-ensembles de la catégorie générale :



<sup>13</sup> On peut lire le numéro que *The Linguistic Review* a consacré à cette problématique (n° 23, 2006).

<sup>14</sup> « Dans une théorie de l'exemplaire, chaque catégorie est représentée dans la mémoire par un grand nuage de tokens mémorisés. Ces mémorisations forment une carte cognitive sur laquelle les souvenirs des instances hautement similaires sont proches les uns des autres et les souvenirs d'instances dissemblables sont éloignés. Quand un nouveau tokens est rencontré, il est catégorisé selon sa proximité avec les exemplaires ».

- Fig. 9 Organisation par exemplaires d'une catégorie

À première vue, le modèle peut apparaître peu réaliste en raison de l'importance donnée à la mémoire d'événements concrets, qui sont, par définition, multiples ; mais sa pertinence en acquisition est de plus en plus mise en évidence, et le modèle tend à expliquer avec succès pourquoi et comment les structures des adultes mêmes, sont incrémentiellement modifiées sur de longues périodes. La linguistique, en revanche, est assez peu armée pour valider ou non cette conception qui, de toute façon, semble le plus souvent « mixée » avec des approches plus abstractionnistes (cf. par exemple [ABB 06]). On soulignera enfin que le linguiste (et musicologue) russe B. Gasparov a proposé dans un ouvrage publié en 2010 une approche de la langue qui nous paraît entièrement compatible avec la théorie des exemplaires : pour Gasparov, la compétence linguistique est entièrement une question de mémoire intertextuelle : tout nouvel énoncé est fondé sur les expériences passées mémorisées par les locuteurs. L'énoncé émerge de fragments différents réarrangés par des « greffes ». On pourrait parler là d'une conception linguistique *radicalement* fondée sur l'usage<sup>15</sup>.

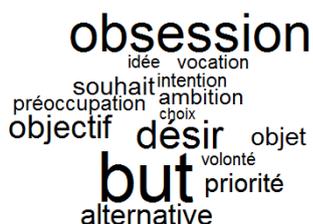
D. Une autre question, qui parfois fait débat en LU, concerne le rapport lexique-construction. Le répertoire étant constitué en réseau, il s'agit de savoir si la relation entre une unité lexicale et les constructions auxquelles il participe est unidirectionnelle ou bidirectionnelle. Autrement dit, puisque le locuteur est en principe sensible au rapport lexème-construction, deux relations peuvent être calculées – ce qui est assez peu souvent le cas. La LU procède généralement à un calcul dit « collostructionnel » [STE 03], calcul qui a été opéré pour la représentation des deux nuages de mots ci-dessus. Il s'agit, compte tenu de la fréquence de la construction, de définir les lexèmes les plus attirés. Par exemple, *objectif* est plus attiré par  $N_{ss}$  *être de Inf.* que *vocation* ou que *souhait* (ce que l'on peut symboliser par  $N_{ss}$  *être de Inf.*  $\leftarrow$  *objectif*). Il s'agit d'un calcul unidirectionnel. Mais un autre calcul est concevable [SCH 10] qui cherche à déterminer quelle est la proportion des emplois d'un lexème dans la construction par rapport à ses autres emplois ; donc,  $N_{ss} \rightarrow N_{ss}$  *être de Inf.* Le premier type de calcul porte sur l'*attraction* du lexème par la construction, le second porte sur la *dépendance* du nom à la construction. On peut à nouveau représenter par des nuages les noms

<sup>15</sup> Sont radicales également du point de vue de l'usage, les perspectives dites fonctionnelles comme celles de S. Thompson (« *what we think of as grammar may be best understood as combinations of reusable fragments* », [THO 02 : 141] et de P. Hopper et sa « Grammaire émergente » (« *The notion of Emergent Grammaris meant to suggest that structure, or regularity, comes out of discourse and is shaped by discourse in an ongoing process* », [HOP 98, p. 156]. Ces perspectives mériteraient de longs développements que nous ne pouvons proposer ici pour des raisons de place.

appréciatifs et bouliques qui consacrent la plus grande part de leurs emplois à la construction :



- Fig. 10 Nuage des noms appréciatifs consacrant une partie de leur emploi à la construction



- Fig. 11 Nuage des noms bouliques consacrant une partie de leur emploi à la construction.

Les résultats sont alors différents que ceux calculés pour l'attraction (fig.7 et 8). Par exemple, *obsession* est peu attiré par la construction, mais lui consacre – à proportion – une partie non négligeable de ses emplois (par comparaison avec les autres noms de modalité boulique).

Ce double savoir, *construction*  $\leftarrow$  *lexique* / *lexique*  $\rightarrow$  *construction*, est-il cognitivement pertinent ? Ces informations quantitatives qui déterminent les degrés d'ancrage, appartiennent-elles à la compétence des locuteurs ? Ici à nouveau, le linguiste ne peut répondre à cette question : il ne peut que mesurer sur corpus les différents rapports entre unités. En revanche, la psycholinguistique est en principe tout à fait armée pour déterminer si ces multiples connaissances sont accessibles aux locuteurs.

#### 11.4. De la linguistique de l'usage à la grammaire du discours

La LU est dite « maximaliste » dans le sens où elle considère qu'aux formes inscrites dans la compétence peuvent être associées des informations de nature non spécifiquement linguistique (dans le sens formel du terme). Une des questions fondamentales selon nous est celle des genres discursifs : si on les voit comme des catégories d'objets (des catégories de textes), ces catégories subordonnent-elles les formes linguistiques à leur organisation ? Autrement dit, les formes « répertoriées » dans la compétence, en plus des informations de fréquences, en plus des informations liées à l'organisation réticulaire, sont-elles également indexées sur les genres ? Depuis la tradition Bakhtinienne, la prise en compte des genres non plus seulement pour des objectifs d'analyse du discours, mais aussi à des fins d'analyse de la langue et de la grammaire, s'est développée de façon spectaculaire. Si la LU a pour objet, sinon de satisfaire un programme substantialiste<sup>16</sup>, en tous cas de renseigner celui-ci par des études empiriques, il est alors évident qu'elle trouve dans les approches « corpus-driven » des alliés sûrs. On peut citer, parmi ces approches, l'école britannique de linguistique de corpus de J. Sinclair<sup>17</sup>. Mais les travaux les plus fondamentaux dans le domaine de l'analyse grammaticale de l'usage en corpus sont sans doute ceux de D. Biber qui ont mis à profit les analyses polyfactorielles pour la description de genres et registres dans un objectif d'élaboration de grammaires de référence telle que la *Longman Grammar of the Spoken and Written English* (Biber et al. 1999). L'intérêt pour la variation syntaxique a également motivé l'analyse grammaticale des genres dont la diversité implique l'existence de formes fonctionnelles associées à des actes communicationnels différents [DOR 10]. Pour le français, J. Deulofeu et J.M. Debaisieux [DEU 12] ont remarquablement formulé un programme pour « une grammaire attentive à la diversité des usages » qui serait, en définitive, la réalisation du souhait exprimé par A. Meillet. Dans une optique descriptive et didactique, D. Siepmann a élaboré un corpus de référence du français et prépare une grammaire (10 volumes !) du français en usage [SIE 15].

Nous voudrions mentionner dans ce qui suit des réflexions maintenant anciennes qui n'ont pas suscité, à l'époque où elles ont été publiées, l'intérêt qu'elles méritent selon nous. Ce sont celles du linguiste J. P. Sueur développées dans un article paru en 1982. L'objectif de l'auteur est de formuler un programme de recherche portant sur la *grammaire du discours* :

« J'appellerai « grammaire du discours » l'étude des régularités dans la distribution, les uns par rapport aux autres, des faits de lexique, de

<sup>16</sup> Au sens de [AUR 92], c'est-à-dire de définir une théorisation linguistique qui ait une correspondance directe avec les fonctionnements psychophysologiques des locuteurs.

<sup>17</sup> Cf. [LEO 08].

---

syntaxe et d'énonciation dans un corpus. Cette grammaire présuppose donc l'étude de régularités lexicales d'une part, syntaxiques et énonciatives d'autre part dans un texte [...]. Autrement dit, le but de cette grammaire est de définir le plus précisément possible les caractéristiques qui permettent d'identifier spontanément un discours : nous reconnaissons le discours de tel individu, de tel groupe, de tel parti, de tel syndicat, etc. Or, si la spécificité du lexique joue un rôle dans ce processus de reconnaissance, d'autres facteurs interviennent : la syntaxe, les faits d'énonciation, mais surtout la connexion entre ces divers types de faits. Tel mot apparaît de manière privilégiée, mais, surtout, il apparaît de manière privilégiée, à telle place et dans tel cadre. Tout cela fait partie de la « compétence » propre au locuteur – et vient se combiner avec les traits qui définissent les diverses formes d'énonciation (une résolution n'est pas un discours ni une réponse à une interview, etc.) » [SUE 82, p. 147-148].

Ce programme a été en partie satisfait par les travaux de Biber à partir de 1988 : les faits lexicaux, syntaxiques et énonciatifs, ainsi d'ailleurs que les corrélations entre ces faits ont été établis sur des genres ou registres différents. L'énonciation possède d'ailleurs un statut particulier dont il nous faut dire deux mots : Bally a tracé la voie à la prise en compte de la subjectivité dans la langue, et Benveniste a définitivement ancré l'énonciation dans la linguistique. Les traces linguistiques de la présence du locuteur dans les énoncés ne sont-ils pas pris en charge par une linguistique de l'usage, c'est-à-dire non pas tant une linguistique de la fréquence et de la norme, que celle de l'activité de langage et d'une pragmatique de la langue ? Cette activité de langage, la place du locuteur, sa relation au monde, l'organisation des données de l'expérience, le point de vue comme condition de la construction du sens, sont des socles communs aux théories de l'énonciation et à la linguistique cognitive de Langacker.

Cependant, le rapport entre compétence des locuteurs et grammaire du discours tel qu'il est formulé par Sueur complique terriblement les modèles d'organisation cognitive proposés par la LU : d'une part, aux genres ne sont pas associées uniquement des informations linguistiques ; des informations sociales leur sont également inhérentes. Par ailleurs, un des points forts de la réflexion de Sueur est de considérer une donnée importante : la répartition non aléatoire du lexique - et on pourrait pousser jusqu'aux structures syntaxiques - dans les différents mouvements / parties d'un texte. Le texte travaillé comme « sac de mots » par la lexicométrie (de l'époque) ne pouvait être analysé comme un tout cohérent, dans lequel opère pourtant une distribution mesurable des unités. Cette dimension « textuelle » des unités (leur localisation en partie prédéterminée dans certains genres) relève de ce

que M. Hoey (2005) nommera bien plus tard la *colligation textuelle*. Beaucoup reste à faire à la linguistique pour identifier cette attribution topologique des unités. Mais là, à nouveau, il s'agirait d'un type de connaissance intégrant la compétence des locuteurs. On peut alors considérer la complexité d'un modèle de la compétence linguistique établi par les théories de l'usage. Comme le dit J.P. Chevrot :

« Ce projet (celui de la LU) souffre de l'incompatibilité des formats dans lesquels sont formulées les connaissances linguistiques et les connaissances indexicales<sup>18</sup>. Plus généralement, nous manquons d'une théorie des liens entre connaissances du monde social et connaissances linguistiques » [CHE 12, p. 57].

### 11.5. Conclusion

Cet article a abordé les questions que ne cessent de poser la perspective fondée sur l'usage au linguiste : la redondance de l'information, peu compatible avec l'élégance des systèmes économiques des approches formelles, les niveaux et les modes de catégorisation (schématisation) des unités linguistiques, la nature même de ces unités (lexèmes, constructions), le rapport statistique des lexèmes aux constructions, l'apport des corpus à l'étude de la cognition linguistique et à la psychologie, les différents types de compétence et leur unification, etc. Ces quelques pages ont pu montrer, que l'usage n'est pas seulement un phénomène pivot indispensable pour appréhender la relation système-norme : elle a aussi des implications en termes de cognition qui nourrissent actuellement de nombreux débats en linguistique. Elle est donc également le lieu où la complémentarité entre empirie et théorie devient indépassable. Nous espérons avoir pu apporter des éléments suffisants au lecteur pour qu'il se forge une opinion de l'intérêt de la LU pour les sciences du langage.

### 11.6. Bibliographie

- [ABB 06] ABBOT-SMITH K. et TOMASELLO M., « Exemplar-learning and schematization in a usage-based account of syntactic acquisition ». *Linguistic Review*, 23, pp. 275-290, 2006.  
 [ADL à paraître] ADLER S. et LEGALLOIS D. « Les noms sous-spécifiés dans le débat parlementaire : analyse collostructionnelle et discursive ». *Langue Française*.  
 [AUR 92] AUROUX S., « La philosophie linguistique d'Antoine Culioli ». In *La théorie d'Antoine Culioli*, Ophrys, pp. 39-60, 1992.  
 [BIB 99] BIBER D. et al., *Longman Grammar of the Spoken and Written English*, Pearson Education ESL, Harlow, 1999.

<sup>18</sup> C'est-à-dire des connaissances sociales, contextuelles, etc.

- [BIB 88] BIBER D., *Variation across speech and writing*. Cambridge University Press, Cambridge, 1988.
- [BLU 12] BLUMENTHAL-DRAMÉ A., *Entrenchment in Usage-Based Theories. What Corpus Data Do and Do Not Reveal About The Mind*, De Gruyter Mouton, Berlin-New York, 2012.
- [BRA 63] BRAINE M., « The ontogeny of English phrase structure: the first phrase ». *Language*, 39, pp. 1-14, 1963.
- [BYB 10] BYBEE J., *Language, Usage and Cognition*, Cambridge University Press, Cambridge, 2010.
- [CHE 12] CHEVROT J.-P., « Que reste-t-il de la distinction compétence-performance dans les théories fondées sur l'usage ? », *Mémoires de la Société Linguistique de Paris*, vol. 21 : pp. 49-68, 2012.
- [CRO 03] CROFT W., *Typology and universals*, (second edition), Cambridge University Press, Cambridge, 2003.
- [DOR 10] DORGELOH H. et WANNER A. (eds.), *Syntactic variation and genre*, De Gruyter Mouton, Berlin-New York, 2010.
- [EGA 08] EGAN Th., *Non-finite complementation: A usage-based study of infinitive and -ing clauses in English*, PaperBack Swap, 2008.
- [FOR 11] FORTIS J., « La notion de grammaire usage-based chez Langacker. Emergence et développement », *Travaux de linguistique*, vol. 62, n° 1, 2011.
- [GAS 10] GASPAROV B., *Speech, Memory, and Meaning Intertextuality in Everyday Language*, De Gruyter Mouton, 2010.
- [GOL 06] GOLDBERG A., *Constructions at Work : The Nature of Generalization in Language*. Oxford University Press, 2006.
- [GRE 06] GREA, P., « La notion d'entrenchment dans le cadre des grammaires cognitives », in D. Legallois et J. François (éds), « Autour des grammaires de constructions et de patterns », *Cahier du Crisco*, 21, 2006.
- [GUI 54] GUIRAUD P., *Les Caractères statistiques du vocabulaire*, Presses Universitaires de France, Paris, 1954.
- [HIL 08] HILPERT M., *Germanic Future Constructions : A Usage-Based Approach to Language Change*, John Benjamins, Amsterdam, 2008.
- [HIN 86] HINTZMAN D., « Schema abstraction in a multiple-trace memory model ». *Psychological Review*, 93(4), pp. 411-428, 1986.
- [HOE 05] HOEY M., *Lexical priming: A new theory of words and language*, Routledge, London, 2005.
- [HOF 14] HOFFMANN Th., *Preposition Placement in English : A Usage-based Approach*, Cambridge University Press, Cambridge, 2014.
- [HOP 98] HOPPER P., « Emergent Grammar », in M. Tomasello (ed.) *The new psychology of language : cognitive and functional approaches to language structure, Volume I*, Psychology Press & Routledge Classic Editions, pp. 155-175.
- [KAB 14] KABATA K. and ONO T., *Usage-based Approaches to Japanese Grammar : Towards the understanding of human language*, John Benjamins, Amsterdam, 2014.
- [LAN 87] LANGACKER R., *Foundations of Cognitive Grammar*. Vol.1. Stanford University Press, 1987.
- [LEG 11] LEGALLOIS D. et FRANÇOIS J., « La Linguistique fondée sur l'usage : parcours critique », *Travaux de linguistique*, vol. 62, no. 1, pp. 7-33, 2011.
- [LEO 08] LEON J., 2008, « Aux sources de la « Corpus Linguistics » : Firth et la London School », *Langages*, vol. 171, no. 3, 12-33.
- [MUK 05] MUKHERJEE J., *English Ditransitive Verbs: Aspects of Theory, Description and a Usage-Based Model*, Rodopi, Amsterdam, 2005.

- [NIN 03] NINIO A., « No verb is an island: Negative evidence on the verb island hypothesis ». *Psychology of Language and Communication*, vol. 7, pp. 3-21, 2003.
- [NOS 92] NOSOFSKY R., « Exemplars, prototypes and similarity rules, in A. F. Healy, S. M. Kosslyn et R. M. Shi rin (eds.), *Essays in honor of William K. Estes: Vol I. From learning theory to connectionist theory*, Erlbaum, Hillsdale, NJ, pp. 149–167, 1992.
- [ORT 16] ORTEGA L., TYLER A., IN PARK H. and UNO M., (éd) *The Usage-based Study of Language Learning and Multilingualism*, Georgetown University Press, 2016.
- [PIE 01] PIERREHUMBERT, J., « Exemplar dynamics: Word frequency, lenition and contrast », in J. Bybee et P. Hopper (eds.), *Frequency and the Emergence of Linguistic Structure*, John Benjamins, Amsterdam/Philadelphia, pp. 137–157, 2001.
- [SCH 10] SCHMID H.J., « Does frequency in text really instantiate entrenchment in the cognitive system? And do we have a quantitative grip on either of them? », in D. Glynn et K. Fischer, *Usage-Based Cognitive Semantics. Corpus-Driven Methods for the Study of Meaning*, Mouton de Gruyter, Berlin, pp. 101-133, 2010.
- [SIE 15] SIEPMANN D., BÜRCEL Ch. et DIWERSY S., « The Corpus de référence du français contemporain (CRFC) as the first genre-diverse mega-corpus of French », *International Journal of Lexicography*, vol. 30, pp. 25-41, 2015.
- [STE 03] STEFANOWITSCH A. et GRIES S., « Collocations : Investigating the interaction of words and constructions », *International Journal of Corpus Linguistics*, vol. 8, pp. 209–243, 2003.
- [SUE 82] SUEUR J.P., « Pour une *grammaire du discours* ». *Mots*, vol. 5, pp. 143–185, 1982.
- [TAY 15] TAYLOR J., *The Mental Corpus: How Language Is Represented in the Mind*, Oxford University Press, 2015.
- [THO 02] THOMPSON S., « Object complements and conversation towards a realistic account », *Studies in Language*, Vol. 26, n° 1, pp. 125-163, 2002.
- [TOM 05] TOMASELLO M., *Constructing a Language: A Usage-Based Theory of Language Acquisition*, Harvard University Press, Cambridge MA, 2005.
- [VER 03] VERHAGEN A. and VAN DE WEIJER J., eds., *Usage-based approaches to Dutch: lexicon, grammar, discourse*, LOT, Utrecht, 2003.
- [WIE 14] WIECHMANN D., *Understanding Relative Clauses: A Usage-Based View on the Processing of Complex Constructions*, De Gruyter Mouton, Berlin – Munich – Boston, 2014.
- [WUL 10] WULFF S., *Rethinking idiomaticity: a usage-based approach*, Bloomsbury Academic, New York, 2010.