

LA PHRASÉOLOGIE
Du figement aux structures lexicales*
Jean-Claude Anscombe
01-2026

Pour citer cette notice :

Anscombe (J.-C.), 2026, « La phraséologie. Du figement aux structures lexicales », in *Encyclopédie Grammaticale du Français*, en ligne : <http://encyclogram.fr>
DOI ; https://nakala.fr/10.34847/nkl.*****

1. À L'ORIGINE ÉTAIENT LES LOCUTIONS

La problématique du figement apparaît avec la notion de *lexique*, laquelle à son tour met en jeu la notion de *mot* : un lexique est une liste de mots, plus ou moins longue selon les objectifs qu'on s'est fixés et les critères choisis. Avant même de nous interroger sur le sens exact de ces deux vocables, notons que leur apparition semble très vraisemblablement liée à celle de l'*écriture*. Le débat sur ce qu'est une écriture n'étant pas le propos de cette notice, nous fixerons arbitrairement les débuts de l'écriture avec l'apparition entre 3200-3000 av. J.C. de l'écriture cunéiforme sumérienne, résultat probable de processus antérieurs de notations pictogrammiques puis idéogrammiques rendues nécessaires pour la gestion d'une société en pleine sédentarisation (Gilmont, 2004: 17 sq.). Les échanges commerciaux avec d'autres peuples ont abouti à la rédaction dès l'époque du Babyloniens Anciens (2000-1600 av. J.C.) d'un nombre important de dictionnaires bilingues sumérien-akkadien (Hollaran, 2006: ii).

Étant donné l'abondance à la même époque de textes littéraires et religieux, les entrées des dictionnaires ne concernaient visiblement pas seulement des mots désignant des objets matériels familiers, mais également des notions abstraites (Hollaran, 2006). Bizarrement, d'autres recueils contenaient des textes de type sapiential, ainsi le proverbe *a-ra-zu a-šed-da šà-ge im-šed-e* ‘une prière est comme de l'eau froide qui tempère l'âtre’ (Alster, 1997 : 8, proverbe n° 141). On peut s'étonner que les recueils de proverbes figurent à part des recueils de mots ‘ordinaires’, puisque à bien y regarder, les deux genres traitent du lexique. On ne voit guère en effet pourquoi les mots composés apparaissent dans les lexiques sumériens, mais non les proverbes, qui sont traités à part. Notons cependant que notre notion moderne de *mot* n'est guère moins étonnante. Tout en clamant en effet que les expressions dites figées (dont les proverbes) sont des entités lexicales, nous construisons des dictionnaires dans lesquels les mots composés et les

* Tous mes remerciements pour leur aide et les nombreuses discussions que nous avons eues sur les thèmes évoqués dans ce texte à : A. Balibar-Mrabti (Prof. émérite, Univ. de Picardie-Jules Verne), D. Flament (Professeur émérite, U. de Paris-Nanterre), Brigitte Lépinette (Prof. émérite, U. de Valencia, Espagne) et I. Tamba-Mecz (Dir. d'Etude émérite, EHESS-CRLAO).

expressions figées ne sont pas des entrées de dictionnaire (et ne sont donc pas des mots), mais figurent sous les mots choisis comme entrées...

On a là un premier exemple – ô combien ancien – d'un phénomène général qui voit dans les mots ‘habituels’ et dans des manifestations comme les proverbes des entités lexicales de natures différentes, sans qu'aucune explication soit fournie qui justifie ce choix. Ce phénomène devait traverser les siècles, à savoir que la présence dans les dictionnaires usuels de ce que nous appellerions aujourd’hui des tournures idiomatiques possède un caractère erratique qui ne laisse pas de surprendre. Certes, on trouve au cours de l’histoire des dictionnaires qui mentionnent de telles tournures : ainsi César Oudin dans son *Tesoro de las dos lenguas española y francesa* (1675), étudié dans Lépinette (1991), mentionne-t-il la tournure *no desplegar su boca* 'ne dire mot, se taire' (s.v. *desplegar* 'déplier, déployer'). Il va même jusqu’à consacrer tout une page à des tournures centrées sur le mot *ojo* 'œil' (s.v. *ojo*). Plus près de nous, Littré (1872 : s.v. *œil*) propose plus de trente locutions formées sur le mot *œil*. Mais ce phénomène n'est pas général. On note en effet que la tendance est à ne pas mentionner ou peu les ‘tournures idiomatiques’¹, quelles qu’elles soient, en particulier dans les dictionnaires de petite et moyenne taille. Elles ne sont cependant pas toujours absentes des ouvrages plus importants. Ainsi le *TLF* mentionne un grand nombre d’expressions formées sur *œil*, mais est beaucoup plus modeste par exemple quant à celles centrées sur *bouche*, puisqu'il ne mentionne pas *ne pas ouvrir la bouche*, non plus que *faire du bouche à bouche* ni *la bouche en cœur*, *provisions de bouche* ou *la bouche pincée*. Sous *main*, on ne trouve pas *avoir la main verte*, *avoir la main sur*, *avoir bien en main* ni *laisser la main*. Ce qui interroge ici n'est pas tant l'absence d'une expression (ou aussi bien sa présence) que les raisons de cette absence (ou de cette présence). Phénomène particulièrement patent dans les manuels pour l'apprentissage de langues étrangères qui, pour chaque thème étudié, fournissent une liste de vocables suivie d'une rubrique où figurent pêle-mêle des locutions verbales, des proverbes, des locutions prépositionnelles, des formules de politesse, des phrases usuelles, etc. De plus, ont été publiés de nombreux recueils ou dictionnaires d'*expressions et locutions* (Rey-Chantreau, 2003), de *proverbes et dictons* (Dournon, 1986), de *locutions idiomatiques françaises* (Lafleur, 1991), pour n'en citer qu'une infime partie. Et, dernier point, il en est ainsi en Europe depuis l'invention de l'imprimerie. Exemple parmi d'autres, l'édition de 1743 du *Dictionnaire de Trévoux* est contemporaine du *Dictionnaire des proverbes françois, et des façons de parler comiques, burlesques et familières* de Panckoucke, paru en 1750 (v. entre autres Badiou-Monferran 2004).

On constate en résumé que, dès l'apparition de recueils de termes, l'homme a généralement séparé ce qu'on pourrait appeler le lexique ordinaire et certaines

¹ Dénomination réactivée récemment probablement sous l'influence de la linguistique américaine. Un des nombreux termes pour désigner ces phénomènes dans les dictionnaires et manuels de langue, parmi *tournures familières*, *expressions*, *expressions toutes faites*, *expressions figées*, *expressions d'usage courant*, *phrases figées*, *phrases toutes faites*, *phrases d'un intérêt courant*, *locutions*, *locutions verbales*, *unités phraséologiques*, *idiomes*, *idiotismes*, etc., et dont l'abondance est symptomatique de l'absence de définition stable du phénomène.

combinaisons dont la nature lui paraissait différente des combinaisons usuelles réalisables à l'aide de ce lexique, en particulier du point de vue de la transparence sémantique. Mais on constate également que, jusqu'à une date récente, aucune caractérisation un tant soit peu méthodique de tels phénomènes n'a été proposée, les auteurs ayant à les illustrer se bornant aux diverses étiquettes que nous avons mentionnées. Or, une étiquette n'a jamais été une explication, ni même l'esquisse d'un concept.

2. LES PRÉCURSEURS

Pendant longtemps donc, ce que nous appelons provisoirement les tournures idiomatiques ont été perçues comme 'différentes', sans qu'on se préoccupe de se demander si elles formaient une catégorie homogène ou bien plusieurs catégories, et encore moins de les caractériser par des propriétés stables, éventuellement linguistiques. Certes, il n'a pas manqué d'auteurs pour signaler cet état de choses, que ce soit le flou terminologique ou le manque de définition. Sur ce dernier point, deux réponses étaient possibles : ou bien les tournures idiomatiques sont des phénomènes linguistiques, identifiables à l'aide de critères spécifiques. Ou bien, à l'inverse, leur origine est de nature fondamentalement extralinguistique². Or, jusqu'à une date récente, c'est la seconde option qui a été retenue de façon très générale. Et ce, entre autres, sur la base d'une hypothèse qu'on rencontre fréquemment en phraséologie : l'équivalence *tournures libres/tournures idiomatiques* et *langue écrite/langue parlée*, qui s'appuie de façon parfois explicite sur une autre équivalence, à savoir *langue culte/langue vulgaire*. Ainsi, Rey-Chantreau (2003 : X) voit dans les locutions :

« des formes convenues, toutes faites, héritées par la tradition ou fraîchement créées, qui comportent une originalité de sens (parfois de forme) par rapport aux règles normales de la langue [...] des expressions [...] le plus souvent imagées, et familières : elles mettent dans le discours une couleur que les énoncés régulièrement produits n'ont pas. En même temps, elles sont fixées, traditionnelles et souvent caractéristiques d'une classe, d'un milieu, d'un état de la société ».

Les locutions sont donc anormales, métaphoriques, familières (d'où leur couleur absente des énoncés 'régulièrement produits'), et socialement marquées. Notons que ce type de définition des locutions verbales est très proche de la définition des tournures argotiques, et le point de vue de Rey-Chantreau est loin d'être isolé. Dans le domaine espagnol, où les tournures idiomatiques (les *modismos* et *frases hechas*) ont donné naissance à une grande quantité de recueils, on parle de *locutions figurées et familières*, de *création populaire* (Pérez-Rioja, 1997, IX), pour lesquelles :

« il est parfois risqué de parler de signification », et qui sont « teintées de l'infinie liberté que procure le populaire, la langue parlée » (Buitrago, 2002, 9 ; notre traduction).

Plus prudents, certains parlent des *idiomatic expressions* comme:

² Cf. Lépinette (2011) pour l'historique des constructions *non naturelles*, ainsi que Fournier (2011) sur un thème proche.

« Those countless expressions which everyone uses, which usually evade the normal rules of grammar, and which often have implications quite unconnected with the normal meaning of the words themselves » (Freeman, 1986, *Preface*).

L'expression est plus modérée, mais le contenu reste le même.

Le début du XX^e siècle voit le renouveau des études sur le langage. Apparaît alors une linguistique qui se veut plus rigoureuse, plus *scientifique*. Il n'est donc pas étonnant que certains esprits plus affûtés que les autres en viennent à se poser la question du statut – linguistique ou non – des tournures idiomatiques.

On trouve une première approche de ce qui sera plus tard appelé *figement* chez Bally (1921 : 66 *sq.*), qui distingue, dans une langue donnée, les *groupements passagers* et les *unités indécomposables*, « entre lesquels se placent des groupes intermédiaires appelés *séries phraséologiques* » (*d°*). Les unités indécomposables, ou *unités phraséologiques* « se reconnaissent à certains indices *extérieurs* et *intérieurs* : les premiers se déduisent de la forme des groupes, les autres (seuls importants) de la manière dont les groupes sont *conçus* par l'esprit [...] c'est-à-dire la manière dont les sujets parlants comprennent les locutions » (*d°*). Selon Bally (= B), les indices extérieurs sont : B₁) le groupe est composé de plusieurs mots séparés par l'écriture ; B₂) ces mots sont disposés dans un ordre invariable, et ne peuvent être séparés par d'autres mots ; B₃) aucun des mots du groupe ne peut être remplacé par d'autres mots. Bally précise d'ailleurs sur la base d'exemples que ces critères ne sont pas suffisants pour caractériser une unité phraséologique. Pour ce qui est des critères intérieurs, ce sont les suivants : B₄) l'équivalence de l'unité examinée avec un mot simple, par exemple *prendre la fuite* et *fuir* ; B₅) l'oubli du sens des éléments, comme dans *avoir maille à partir* ; B₆) la présence à l'intérieur de l'unité d'un archaïsme. Pour Bally, un archaïsme est un composant d'une unité qui pris isolément, n'est pas compris du sujet parlant, et ne le devient que parce qu'inclus dans une unité. C'est le cas de *foi* dans *de mauvaise foi*. De façon assez inattendue, Bally affirme aussi l'existence d'archaïsmes de syntaxe, par exemple l'inversion verbe-objet dans *il faut raison garder*, dans lesquels il voit des sortes de 'moules' permettant de créer des expressions nouvelles, sans voir la possible interférence entre *archaïsme* tout court qui est sémantique, et *archaïsme de syntaxe*, qui ne l'est pas. En conclusion, Bally inscrit l'étude des unités phraséologiques à l'intérieur de celle de « tout fait de langage qui offre un sens complet et autonome, c'est-à-dire qui correspond à une unité de pensée », les *unités lexicologiques*. On en conclut immédiatement que pour Bally, les unités phraséologiques relèvent du lexique.

On excusera ce long développement consacré au point de vue de Bally sur les tournures idiomatiques. Outre son intérêt historique, son approche présente un certain nombre de points communs avec des approches plus récentes, nous le verrons en détail plus avant. Nous noterons simplement que B₁ est en fait l'hypothèse de la *polylexicalité* des tournures idiomatiques, redécouverte dans Greciano (1983).

Quelque temps plus tard, Jespersen (1924 : 14-22) propose de distinguer les *expressions figées* (*formulas*) et les *expressions libres* (*free expressions*). Il remarque en effet qu'en

langue, il y a des 'choses' qui ont un caractère formulaire, i.e. auxquelles on ne peut apporter aucun changement (*op.cit.*, p. 8):

«... Some things in language-in any language-are of the formula character; that is to say, no one can change anything in them ... ».

Et d'illustrer ce propos par des cas comme *How do you do ?, Thank you, Good morning*, pour lesquels il note que tout y est fixe, et qu'on ne peut même pas en modifier la ligne mélodique. De plus, le sens du tout n'est pas déductible du sens des mots qui composent l'expression considérée. On notera que ces critères sont peu ou prou les critères B₁, B₂, B₃ et B₅, auxquels Jespersen ajoute celui de l'invariabilité prosodique, Bally conservant en revanche la paternité de B₄ et B₆. Pour Jespersen, les expressions figées sont apprises, comprises et utilisées en bloc, alors que l'usage des expressions libres suppose à chaque fois une 'activité mentale' de créativité absente dans le cas des expressions figées ou formules. Bien sûr, cette séparation est moins idyllique qu'il n'y paraît : parler ne consiste pas seulement à combiner des mots du dictionnaire selon les règles de la grammaire en vigueur, mais également à maîtriser certains 'blocs'. Même les constructions libres peuvent résulter d'habitudes de langage, de tendances à reproduire ce qui a déjà été produit ou entendu, ce qui atténue la différence entre expressions libres et figées. À quoi on pourrait ajouter que certaines expressions qu'on aurait tendance à voir comme figées (au sens de Jespersen) admettent certaines variations, ainsi (Anscombe, 2008) : *pâle comme (un linge + un mort + la mort + une feuille de papier + une endive)* ; ainsi que l'introduction de certains éléments de type souvent adverbial: *en tomber sur le derrière / en tomber carrément sur le derrière, être fait pour être x comme moi pour être pape / être fait pour être x à peu près comme moi pour être pape*, parmi beaucoup d'autres cas.

On pourrait dire, pour résumer ce qui précède, que tant Bally que Jespersen sont à l'origine de la notion de *figement* en tant que phénomène linguistique, phénomène qu'ils tentent de caractériser par diverses propriétés : syntaxiques (les expressions figées ont une syntaxe fixe), sémantiques (le sens des expressions figées n'est pas déductible du sens de leurs composants), lexicales (on ne peut remplacer les unités lexicales d'une expression figée par d'autres unités, même synonymes), prosodiques enfin (on constate la présence d'intonations spécifiques). Tous deux constatent également l'existence de 'cas intermédiaires' entre expressions figées et expressions libres, qui rend difficile l'identification de l'appartenance d'une combinaison à l'une ou l'autre catégorie.

Bizarrement, et jusqu'aux années 1970-80, ces textes, qui ouvraient un champ encore inexploré de la linguistique, sont restés sans influence sur l'orientation des recherches de cette discipline. Si l'on peut trouver ça et là quelques travaux touchant à ces problèmes, il faudra cependant attendre les années 1970-1980 pour voir apparaître un véritable courant de recherches autour de cette thématique, en particulier sous l'impulsion de la grammaire générative.

3. L'APPORT GÉNÉRATIVISTE À LA NOTION DE FIGEMENT

3.1. *Le figement dans la grammaire générative transformationnelle*

C'est en effet selon nous à la grammaire générative transformationnelle qu'on doit la première formulation générale explicite des phénomènes de figement comme une catégorie linguistique à part entière, c'est-à-dire non comme des exceptions plus ou moins étranges, mais comme des phénomènes linguistiques devant être engendrés dans le système linguistique adopté pour traiter la langue dans son ensemble. Nous nous permettrons d'insister sur ce point, car il représente une étape décisive dans l'histoire des phénomènes de figement. Décider que le figement définit une catégorie linguistique implique par là-même qu'il ne peut recevoir un traitement adéquat que dans et par rapport à une théorie de la langue qui définit précisément de telles catégories.

Nous commencerons par rappeler brièvement les principes de base de la grammaire transformationnelle, historiquement la première étape de la grammaire générative³. L'idée de base est de fonder la description d'une langue sur un système formel, c'est-à-dire un ensemble d'axiomes de départ, et de règles et d'instructions dont l'application mécanique engendrera les énoncés admissibles d'une langue, et eux seuls. Dans la *théorie standard*, une grammaire transformationnelle comprend trois composants : un composant *syntaxique*, lui-même divisé en un composant *syntagmatique* et un composant *transformationnel* ; un composant *sémantique*, et un composant *phonologique*. Au départ, le composant syntagmatique opère sur des suites de symboles, et sa fonction consiste à appliquer des règles de réécriture aboutissant à une nouvelle séquence formelle composée de symboles. D'autres règles, les *règles d'insertion lexicale*, font correspondre à certains symboles de cette séquence, les symboles dits terminaux, de véritables entités linguistiques, mots ou morphèmes. Le résultat obtenu, ou *structure profonde*, passe ensuite par le composant transformationnel, chargé de déplacer, supprimer ou ajouter des éléments. Une des transformations les plus célèbres était la transformation passive : elle dérivait une phrase passive de la forme active correspondante. Le résultat, ou *structure superficielle*, traversait le composant phonologique qui lui associait une *représentation phonétique*. Dernier point, et non des moindres : l'interprétation sémantique, due au *composant sémantique*, intervenait au niveau de la structure profonde. C'est ainsi qu'une phrase active et son éventuelle transformée passive avaient, dans ce modèle, la même représentation sémantique. Voici, en très simplifié, l'essentiel du fonctionnement d'une grammaire transformationnelle standard, suffisant pour appréhender les problèmes posés par des tournures comme *to kick the bucket* 'casser sa pipe', *to spill the beans* 'vendre la mèche', *to break the ice* 'rompre la glace', *to beat around the bush* 'tourner autour du pot', etc., pour reprendre des exemples restés célèbres.

³ On pourra se reporter pour plus d'information à Chomsky (1965), Chomsky (1970 : 60-110), et pour une étude très complète à Ruwet (1967 : V). Le texte ci-dessus ne prétend pas être un exposé complet sur la grammaire transformationnelle, mais le rappel du cadre théorique qu'elle propose dans la mesure où il influe sur le traitement des phénomènes de figement.

Ce modèle transformationnel a suscité passions et critiques. Son principal mérite est d'avoir proposé un *modèle scientifique* explicite et de s'y être tenu, en en permettant du même coup la falsifiabilité. On peut en revanche ne pas partager certaines des options théoriques qui y sont proposées. Là n'est pas notre propos. Indépendamment donc de la validité théorique du modèle générativiste évoqué ici, le point qui nous intéressera est son incidence sur la conceptualisation de la notion de figement qu'il induira. Dès les premiers textes génératifs en effet, c'est-à-dire dès la fin des années 1960, les tenants de cette ligne théorique se sont aperçus que certaines combinaisons ne pouvaient pas être traitées telles quelles dans ce modèle, et ont ainsi découvert le phénomène des *idioms*⁴, largement utilisé par la suite pour critiquer entre autres choses le traitement de la sémantique dans un tel cadre. Les raisons invoquées pour distinguer une telle classe étaient au nombre de trois :

(a) *La non-actualisation des éléments (figement référentiel)* :

Improprement appelé *figement référentiel*, ce phénomène concerne la non-actualisation des éléments lexicaux qui se manifeste en particulier à travers l'impossibilité de modifier le(s) déterminant(s). Cette propriété est en fait semblable à B₃ évoquée plus haut. Elle a donc été plus particulièrement étudiée dans le cas des éléments nominaux d'une séquence dont il s'agit d'apprécier le caractère plus ou moins figé. Dans le cadre générativiste de départ, le raisonnement est nécessairement mené au niveau de la structure profonde. Reprenons alors le cas antérieurement mentionné de *casser sa pipe* (*To kick the bucket* dans les textes anglais), unanimement considéré comme un cas-type de figement. Une telle suite polylexicale devrait être engendrée en structure profonde à partir d'une suite symbolique de type GV + Dét + N⁵, qui fournirait par insertion lexicale des symboles terminaux *casser*, *son* et *pipe*, la suite étudiée *casser sa pipe*. *Casser sa pipe* serait ainsi engendrée de la même façon que *briser sa pipe*, *casser son tuyau*, *massacer son tuba*, etc. En d'autres termes, *casser sa pipe* appartiendrait à un ensemble de constructions apparentées où les éléments de chaque catégorie seraient à choisir dans un paradigme, et serait alors considérée comme une *structure libre*, ce qui n'est pas le cas. À la possibilité de par exemple *un plongeur a massacré (mon + ton + son) tuba*, on opposera les restrictions *un plongeur a cassé (*ma + *ta + sa) pipe*, *tu as cassé (*ma + ta + *sa) pipe*, etc. L'usage de *casser sa pipe* 'mourir' implique entre autres, à l'inverse des structures libres, une coréférence avec le sujet de la phrase où elle apparaît, en l'occurrence *le plongeur*. On voit ainsi, à partir de cet exemple ultra simplifié, comment la grammaire transformationnelle se voit interdire l'engendrement de *casser sa pipe* en structure profonde.

⁴ Par exemple, on trouve déjà chez Katz & Postal (1963), l'idée que les *idioms* peuvent fonctionner comme des unités lexicales, et qu'ils ne satisfont pas à la règle de compositionnalité. Cet article distingue également entre *lexical idioms* (noms composés, adjetifs, etc.) et *phrase idioms* (essentiellement verbe + objet).

⁵ Nous laissons volontairement de côté tous les aspects techniques. Ils ne changent rien au problème, ni au raisonnement.

(b) *Le blocage de la possibilité de certaines manipulations (figement transformationnel)* :

C'est un des critères les plus connus, ce qui est sans doute dû au fait que c'est aussi le plus facile à mettre en évidence. Il s'agit de la réticence des suites polylexicales figées à admettre les modifications, ou du moins certaines d'entre elles. Ce critère est proche du B₂ de Bally. Dans le cadre générativiste initial en effet, les modifications éventuelles sont opérées par le composant transformationnel, et ne peuvent donc être que des transformations disponibles dans ce composant. Le cas plus célèbre est la transformation de passivation, illustrée par les exemples suivants :

- (1) (a) Max a pris le mors aux dents.
(b) *Le mors aux dents a été pris par Max.
- (2) (a) John broke the ice.
(b) ??The ice was broken by John.

Que l'on opposera à :

- (3) (a) Max a photographié la scène de crime ;
(b) La scène de crime a été photographiée par Max⁶.

(c) *L'opacité sémantique (figement sémantique)* :

Il avait déjà été remarqué et fréquemment signalé que le sens global d'une expression (supposément) figée n'est en général pas déductible du sens des éléments qui la composent formellement : c'est le *principe de non compositionnalité* des expressions figées (v. Tamba 2014a, 2014b), déjà présent dans le B₅ de Bally. Ainsi, le caractère familier des mots *mettre*, *dans*, *doigt* et *œil*, qui relèvent pourtant d'un registre courant de la langue, ne permet pas d'en déduire la signification de la locution *se mettre le doigt dans l'œil*, à savoir 'se tromper lourdement'. Et il ne s'agit pas là d'un cas isolé, au vu d'autres locutions comme *entre chien et loup*, 'au crépuscule', ou encore : *au doigt et à l'œil*, *cousu de fil blanc*, *effeuiller la marguerite*, *enfoncer une porte ouverte*, *en plein dans le mille*, *faire chou blanc*, *la langue de bois*, *pédaler dans la choucroute*, *prendre les patins* (de qqn), *un de ces quatre matins*, etc., pour ne citer qu'une (toute petite) partie des locutions du français contemporain, qu'on estime à plusieurs milliers.

On remarquera que la non-actualisation des éléments correspond au fait que, dans le modèle transformationnel, les suites dites figées ne sont pas directement engendrables au niveau du composant syntagmatique. Cette possibilité supposerait en effet que chaque élément de la suite appartienne à un paradigme indépendant, ce qui n'est pas le cas, nous l'avons vu. La locution *casser sa pipe* n'est pas engendrée comme *briser sa pipe*, cette dernière étant engendrée en revanche comme *briser sa bouffarde* ou encore *casser son fume-cigarette*. De plus, ces locutions figées commutent fréquemment avec des unités

⁶ Notons que ces critères ne sont pas toujours fiables. La passivation peut être difficile pour des raisons syntaxiques n'ayant rien à voir avec le figement. Ainsi dans le cas de *Lia_i a vendu sa maison_i* / ??*Sa_i maison a été vendue par Lia_i*, où le problème signalé tient à une propriété du possessif, et non à un quelconque figement.

lexicales simples, ainsi : *casser sa pipe/mourir, prendre son pied/s'éclater, cordon bleu/chef, prendre les patins/soutenir, faire chou blanc/échouer, pédaler dans la choucroute/patauger, ruer dans les branards/se révolter, battre la campagne/divaguer, faire étalage de/exhiber, dans les grandes largeurs/complètement, bas de plafond/borné*, etc. D'où l'idée avancée dès le début de l'application de la théorie transformationnelle à des phénomènes linguistiques, et qui devait être remaniée par la suite, d'introduire de telles locutions 'en bloc', ce qui ne peut être fait qu'au niveau lexical. Dans un tel cadre, les locutions dites figées sont donc naturellement considérées comme faisant partie du lexique, au même titre qu'un mot 'simple'. Le blocage éventuel de certaines manipulations renvoie au composant transformationnel qui agit au niveau de la suite considérée, et non à l'intérieur d'une unité lexicale de cette suite ; et enfin, le critère de non-compositionnalité, le seul à faire à peu près l'unanimité⁷, correspond bien sûr au composant sémantique.

On peut donc dire qu'à l'intérieur de ce cadre théorique de la grammaire générative transformationnelle : a) Le figement reçoit une définition : est figée une suite d'éléments qui ne peut être engendrée, au contraire des structures libres, par le composant syntaxique ; b) Les critères permettant d'identifier une expression figée correspondent à des irrégularités relevant des composants de la grammaire générative transformationnelle, i.e. des trois étapes du mécanisme.

3.2. Le figement dans les grammaires transformationnelles : limites et problèmes

C'est dans le cadre générativiste même que ce traitement des phénomènes de figement a suscité de nombreuses critiques et remaniements, en particulier pour ce qui est des critères évoqués en a) et b). Nous nous limiterons aux principaux arguments évoqués par les adversaires non du figement, mais de cette conception du figement spécifique à la grammaire générative transformationnelle. D'une part, les travaux sur le figement sont extrêmement abondants, et il est impossible de tous les présenter dans un espace raisonnable. En fait, l'essentiel des débats sur le figement, ses caractéristiques et sa place dans une théorie linguistique se limitent, jusqu'aux années quatre-vingt, à un duel entre grammaire générative transformationnelle et sémantique générative. Dès 1967 en effet, et à la suite des travaux principalement de Fillmore (1968), Lakoff (1968), Mc Cawley (1968a, 1968b), Ross (1967), entre autres, est apparue une rivale de la grammaire chomskienne défendue par Chomsky et Katz & Postal, à savoir la *Sémantique générative*. Tout se joue autour des places respectives de la syntaxe et de la sémantique (et aussi des rapports qu'elles entretiennent). De façon quelque peu caricaturale, on peut dire que la grammaire générative transformationnelle engendre les structures *de surface* au travers d'une structures *profonde* de nature syntaxique, alors que la sémantique postule

⁷ Moyennant bien souvent cependant l'hypothèse d'une possible gradation de cette non-compositionnalité, sans que cette gradation fasse l'objet de définitions, et encore moins de critères de repérage.

l'inverse : la base est une structure profonde de nature sémantique⁸. Ces deux théories utilisant l'idée de représentation profonde abstraite, on voit immédiatement l'importance que joue dans ce choix la place de l'insertion lexicale. En effet, l'insertion lexicale d'expressions figées pose des problèmes à la fois de syntaxe (impossibilité de certaines transformations) et de sémantique (par exemple la non-satisfaction du principe de compositionnalité). Il n'est donc pas étonnant que la discussion entre ces deux théories ne concerne pas directement la définition du concept de figement, mais le poids des phénomènes de figement dans le choix entre syntaxe générative et sémantique générative, c'est-à-dire entre une représentation profonde de nature syntaxique opposée à une représentation profonde de nature sémantique. Ce que nous illustrerons brièvement par le traitement proposé par Fraser (1970) des phénomènes de non-compositionnalité, ce trait lui servant de définition du concept.

Dès Fraser (1970 : 22) en effet, première étude des *idioms* dans un cadre chomskyen, le figement est rattaché aux *idioms* que Fraser définit de la façon suivante :

I shall regard an idiom as a constituent or series of constituents for which the semantic interpretation is not a compositional function of the formatives of which it is composed. Expressions such as *figure out*, *make love to*, *beat around the bush*, *by accident*, *pass the buck*, and *has the cat got your tongue* are all cases of what I would argue are idioms.

Se plaçant d'entrée dans le cadre de la grammaire transformationnelle et prolongeant Katz & Postal (1963), Fraser s'attaque au problème de la non-compositionnalité des différents types d'idiomes qu'il distingue : mono-morphématiques (*throw*), polymorphématiques (*lexical idioms*, ainsi *overturn* 'renverser'), ou encore polylexicaux (*phrasal idioms*, ainsi *Has the cat got your tongue ?* 'T'as perdu ta langue ?'). Ils posent en effet deux questions principales, qui seront au cœur d'un débat qui devait durer plusieurs années : a) Comment représenter la signification de tels idiomes dans une grammaire transformationnelle ; et b) Comment rendre compte de l'incompatibilité entre ces *phrasal idioms* et certaines transformations syntaxiques, comme la passivation.

Pour répondre à ces deux questions, Fraser imagine un mécanisme complexe de *Hierarchie de congélation* ('Frozenness Hierarchy') à cinq niveaux pour chaque type de transformation, tout idiom se voyant assumer un 'degré' de congélation. La représentation formelle d'un idiom dans le lexique comprendra alors obligatoirement l'indication de ce degré de congélation pour une transformation donnée. La seule différence entre un idiom et sa contrepartie 'littérale' résidera précisément dans la valeur de ce degré de figement, de *congélation*, qui va de 0 (complètement congelé) à 5 (congélation minimale) pour chaque type de transformation. Tous les efforts de Fraser (1970) consistent donc à fournir une représentation qui ne remette pas en cause le modèle transformationnel, tout en ne faisant pas des *idioms* quels qu'ils soient des exceptions à traiter à part. Il maintient ce faisant l'engendrement syntaxique de base et le mécanisme d'insertion lexicale en aménageant une représentation lexicale – une sorte de 'passeport

⁸ Nous laissons de côté la question tout aussi fondamentale de la nécessité de telles structures profondes. On trouvera l'essentiel de la discussion sur ces différents points dans Dubois-Charlier (1972) et surtout Galmiche (1972).

transformationnel' valable à la fois pour les *idioms* et les autres structures syntaxiques, quel que soit leur type.

Cette solution, qui sera ultérieurement (et diversement) reprise sous l'étiquette de 'degré de figement', a été fortement critiquée, ne serait-ce que parce qu'elle fait abstraction des diverses possibilités de relation entre un idiome et ses constituants, c'est-à-dire des types et degrés de compositionnalité possibles, déterminants pour la compréhension et probablement aussi pour l'apprehension de la nature syntaxique de cet idiome (Wood 1986 : 19). Fraser remarquait à ce propos que les impossibilités transformationnelles peuvent varier d'un idiome à l'autre, comme illustré par :

- (4) (a) Max s'est plaint qu'on lui avait mis des bâtons dans les roues.
(b) Max s'est plaint que des bâtons lui avaient été mis dans les roues⁹.
- (5) (a) Lia lui a tenu le crachoir pendant une bonne heure ;
(b) ?Pendant une bonne heure, le crachoir lui a été tenu par Lia.

De tels faits ne sont nullement exceptionnels : Ruwet (1983, 1991 : 171-251) montre en outre que, s'il existe bien un large éventail de contraintes relatives aux transformations syntaxiques que peuvent subir des idiomes, elles ne sont a priori pas prévisibles pour un idiome donné.

C'est en particulier à Nunberg, Sag & Wasow (1994) que l'on doit vraisemblablement l'examen le plus détaillé à la fois sous l'angle de l'opposition syntaxe/sémantique (du moins dans le cadre de la grammaire générative), et d'autre part, sous celui de la définition du concept d'*idiom*. Comme le déclarent d'entrée les auteurs (491-92):

In this paper, we examine various dimensions of idiomaticity and their relation to grammatical theory. Our conclusion will be that idioms provide no evidence bearing one way or the other on such syntactic issues. As we will show, there are compelling reasons to believe that the majority of phrasal idioms are in fact semantically compositional, and that the very phenomenon of idiomaticity is fundamentally semantic in nature. Much of the literature on the syntax of idioms is thus based on the misconception that no such semantic compositionality exists. Rather than providing evidence for theories of transformations, phrase structure, grammatical functions, thematic roles or otherwise unnecessary grammatical devices (e.g. 'quasi roles'), the various constraints on idioms that have been observed are better explained in consequence of either semantic properties or else broader tendencies in the figurative use of language.

D'autant plus que (*idem*, 492):

Attempts to provide categorical, single-criterion definitions of idioms are always to some degree misleading and after the fact. In actual linguistic discourse and lexicographical practice, 'idiom' is applied to a fuzzy category defined on the one hand by ostension of prototypical examples like English *kick the bucket*, *take care of NP*, or

⁹ En voici un exemple attesté (non unique) pour les sceptiques : « ...juger les coupables si longtemps après leurs crimes comportait ce risque, c'est sans doute pour cela que **tant de bâtons ont été mis dans les roues** de ceux qui voulaient organiser le procès des Khmers rouges responsables du génocide ... » (*SketchEngineTenTen23*)

keep tabs on NP, and on the other by implicit opposition to related categories like formulae, fixed phrases, collocations, clichés, sayings, proverbs and allusions-terms which, like 'idiom' itself, inhabit the ungoverned country between lay metalanguage and the theoretical terminology of linguistics.

En particulier, dans le cadre de la grammaire générative, l'idiomaticité a très largement été assimilée dès le début, à la non-compositionnalité. Et, selon Nunberg, Sag & Wasow, une telle vision d'homogénéité du phénomène passe à côté d'un certain nombre de caractéristiques de l'idiomaticité, et confond entre autres la conventionnalité (*conventionality*) et le sens figuré (*figuration*). La conventionnalité désigne la propriété que possède le sens des idiomes de ne pouvoir être totalement prédictible à partir du sens qu'ont leurs composants lorsqu'ils figurent dans des combinaisons libres. Le sens figuré en revanche renvoie à la particularité sémantique de certains idiomes de faire jouer une métaphore, ainsi le fameux exemple *to kick the bucket* 'mourir', l'équivalent anglais du français 'casser sa pipe'. Si les sujets parlants ne savent pas toujours expliquer le lien entre le sens littéral et le sens métaphorique, ils sont cependant presque toujours capables de percevoir la présence d'un tel sens métaphorique, et même d'assigner un sens littéral à l'idiome envisagé.

En résumé, Nunberg, Sag & Wasow proposent donc de distinguer d'une part, les expressions idiomatiques combinatoires (*idiomatically combining expressions*), dont le sens, bien que conventionnel, ainsi *to spill the beans* ('vendre la mèche'), est distribué sur les différentes composants de l'expression. Et les syntagmes idiomatiques (*idiomatic phrases*), qui ne montrent pas une telle répartition, par exemple *to kick the bucket*, ou encore *to eat one's heart out* ('être inconsolable'). Cette distinction, qui est sémantique, présente l'avantage de rendre compte de certaines irrégularités syntaxiques considérées comme typiques des idiomes alors qu'elles ne le sont que d'une sous-catégorie, les expressions ou syntagmes idiomatiques. On explique ainsi aisément le comportement apparemment erratique du fameux critère de passivation prétendument typique de la conventionnalité, alors qu'il ne l'est que du sous-ensemble des syntagmes idiomatiques, comme on le constate sur les exemples suivants des auteurs eux-mêmes :

- (6) (a) *The beans were spilled by Pat.*
(b) **The bucket was kicked by Pat.*

On notera que *to spill the beans* et *to kick the bucket* ont des formes superficielles très proches. Cependant, *to spill the beans* 'divulguer l'information' laisse rapprocher *to spill* de *to divulge*, et *the beans* de *the information*, ce qui en fait une expression idiomatique combinatoire, opération qui n'est pas possible avec *to kick the buck*¹⁰. La différence entre les deux expressions tient au fait qu'elles appartiennent à des sous- catégories sémantiques différentes : toutes deux sont conventionnelles, mais les conventions peuvent être soit relatives aux composants séparément (*to spill the beans*), soit à leur collocation

¹⁰ Bien entendu, nous nous contentons ici de résumer le fil conducteur du raisonnement. On se reporterà au texte pour les détails de l'argumentation. Fait amusant : *vendre la mèche* et *casser sa pipe* présentent pour le français la même opposition : *la mèche a été vendue/*sa pipe a été cassée*.

(*to kick the bucket*). Dans *to spill the beans*, une convention relie *to spill* à *to divulge*, et une autre convention *the beans* à *the information*. Dans *to kick the bucket*, c'est la collocation elle-même des éléments qui est l'objet d'une convention¹¹. D'où la différence de comportement. Ce phénomène étant sémantique, on en déduit que le comportement syntaxique de tels idiomes est dérivé de leur comportement sémantique, et non l'inverse. Et que donc envisager la structure de base comme syntaxique – et non sémantique – ne pourra rendre compte des phénomènes.

En résumé, on doit reconnaître avec Nunberg, Sag & Wasow, que la confusion entre *non-compositionnel* et *conventionnel* conduit à considérer comme de même nature sémantique des phénomènes qui ne le sont pas. Certes, si l'on part du principe que ce qui caractérise les idiomes est la non-compositionnalité, et que cette non-compositionnalité est due au caractère conventionnel de la composition de ces idiomes, on aboutit à l'inévitable conclusion que les idiomes sont conventionnels. Mais l'erreur de raisonnement, nous l'avons vu, est de croire que cette conventionnalité se manifeste toujours de la même façon. Or, ce que montrent Nunberg, Sag & Wasow, nous l'avons vu, est qu'elle peut en fait agir soit séparément au niveau d'un ou plusieurs composants, soit au niveau cette fois de la collocation des différents composants de l'expression. Et cette différence définit deux catégories distinctes d'idiomes, montrant des comportements ‘transformationnels’ différents. D'où l'idée que les propriétés des idiomes relèvent basiquement de la sémantique, et ne peuvent être expliquées dans une théorie syntaxique chomskyenne. On peut cependant regretter que nos auteurs soient essentiellement préoccupés de choisir entre une linguistique à base sémantique face à une linguistique à base syntaxique, et se contentent de considérer les idiomes comme étant fondamentalement non-compositionnels, sans guère fournir de critères opératoires pour en faire une catégorie repérable, ni se poser la question de l'éventuelle hétérogénéité de la catégorie, voire de l'existence de sous-catégories distinctes quant à leurs propriétés. En d'autres termes, les phénomènes de figement sont vus par Nunberg, Sag & Wasow comme des observables externes à la théorie proposée, qu'elle soit syntaxique ou sémantique. Or, les notions de non-compositionnalité et de conventionnalité qu'ils présentent sont quelque peu floues, et la non-prédictibilité du comportement et syntaxique et sémantique des idiomes étudiés, visible sur la dispersion des propriétés constatées, laisse prévoir une complexité qui va bien au-delà de la simple intuition de l'existence d'un phénomène général de figement. Elle renvoie en fait à une théorie de la formation lexicale qui reste à construire, du moins à l'époque de Nunberg, Sag & Wasow (1994), et où les phénomènes de figement trouveraient leur place.

4. L'ÉTAT DES LIEUX POST GÉNÉRATIVISTE

Nous venons de voir que le figement n'est pas étudié dans le cadre générativiste en tant que phénomène, mais surtout en tant que source d'arguments pour faire pencher la théorie

¹¹ Ce critère est proche du B₄ de Bally. Il semblerait en effet que les syntagmes idiomatiques comme *to kick the bucket* soit équivalents à des mots simples – ici *to die*, en raison de la collocation existante entre les composants, et non les expressions idiomatiques combinatoires, comme *to spill the beans*. Ce point reste à étudier en détail.

générativiste vers telle ou telle autre conception. Elle aura cependant eu le mérite indiscutable de faire connaître ce (vaste) champ à une linguistique à l'époque en pleine effervescence, et à laquelle elle proposait une approche originale sur deux points. Le premier point est l'exigence clairement formulée de se donner un cadre scientifique pour l'étude des langues, et de s'y tenir. On peut certes être en désaccord avec telle ou telle option retenue, mais on ne peut dénier à la grammaire générative le strict respect de cette contrainte. On peut la rapprocher dans son esprit – toutes proportions gardées - de la tentative de Russell & Whitehead d'inventer un langage (la logique) assurant la scientificité du raisonnement mathématique. Un second point est l'irruption en linguistique par le biais de la grammaire générative – et en particulier en sémantique – d'une linguistique de la phrase là où n'existaient avant guère autre chose qu'une sémantique lexicale¹². Il s'agit là d'un véritable changement de paradigme au sens de Kuhn (1962), et qui devait avoir un impact indéniable sur l'orientation des études linguistiques, y compris lexicales. On passe du mot ou de la locution comme entité lexicale au mot comme composant phrasique.

4.1. Les travaux de Maurice Gross et de son école : le lexique-grammaire

Il faudra attendre les travaux de M. Gross et de son équipe le LADL (Laboratoire d'Automatique Documentaire et Linguistique), pour voir apparaître en France une théorie linguistique du lexique – le *lexique-grammaire* - qui ne soit pas ... lexicaliste, i.e. qui ne considère pas qu'indépendamment de ses propriétés syntaxiques, un mot possède un *sens*, c'est-à-dire une valeur sémantique propre, et qu'il suffit alors de combiner le sens des mots d'un texte pour avoir accès au sens de tout ou partie du texte¹³. Pour M. Gross en fait, très influencé par les travaux de Zellig H. Harris, la distinction entre les propriétés syntaxiques de deux occurrences d'un même mot et les différences interprétatives (sémantiques donc) de ces deux occurrences sont liées. Configurations syntaxiques et configurations sémantiques d'un mot (au sens habituel) ne sont donc pas autonomes, et on ne saurait étudier l'une sans l'autre. Nous illustrerons cette remarque sur un exemple repris de G. Gross (2012 : 12), celui du verbe *abattre*. Tout francophone est vraisemblablement persuadé de pouvoir en décrire aisément le sens. Pourtant, un examen plus poussé montre des différences selon l'environnement. On peut ainsi dire sans problème *abattre un arbre*, *abattre un homme*, *abattre un avion*, entre autres. Mais les problèmes commencent lorsqu'on tente de fournir des synonymes verbaux : *couper* (*un arbre* + **un homme* + **un avion*), *descendre* (? *un arbre* + *un homme* + *un avion*), *tuer* (??*un arbre* + *un homme* + **un avion*), etc. Ces divergences de comportement indiquent visiblement des différences sémantiques, lesquelles s'accompagnent de différences syntaxiques et morphologiques. Ainsi, la nominalisation de *abattre*, possible avec *arbre*, i.e. *l'abattage d'un arbre*, est impossible dans les autres cas envisagés : **l'abattage* (*d'un homme* + *d'un avion*). Et au sens psychologique de *abattre* – celui qui correspond à *Abattu par son échec, il démissionna de son poste*, la nominalisation correspondante est

¹² Quand Bréal, en 1897, définit la sémantique comme *la science des significations*, il parle de la signification des mots.

¹³ Notons que cette hypothèse que refuse M. Gross est faite par la grammaire transformationnelle de Chomsky, mais aussi en grande partie par les grammaires scolaires.

abattement. La nominalisation en *-age*, très productive en français contemporain dans le domaine technique, est en fait typique des verbes d'action. Alors que la nominalisation en *-ment* est fréquente dans une sous-classe de ‘noms de sentiment’, ceux qui réfèrent à la réaction à un stimulus externe : outre *abattement*, citons *agacement*, *apeurement*, *étonnement*, *pressentiment*, *ressentiment*, etc.

De cette approche de M. Gross découle l’idée qu’un texte, quel qu’il soit, y compris limité à une simple phrase (au sens grammatical habituel), ne se résume pas à une simple combinaison de mots, mais qu’il fait intervenir des *structures élémentaires* liées aux mots, appelées *schémas prédicatifs*, qu’il convient donc d’identifier lors de l’analyse du texte. Un schéma prédicatif est formé d’un *prédictat* et de ses *arguments*, en un sens que nous allons sommairement définir, et représente ce qu’on appelle une *phrase simple*, dans la mesure où elle ne comprend aucune sous-structure analysable à son tour comme un schéma prédicatif (Boons-Guillet-Leclère 1976 : 33) :

... on entend par *phrase simple* une phrase dont le verbe ne prend pas de complétive ou d’infinitive comme complément. Ceci revient à appeler *phrase simple* une phrase à un seul verbe, si l’on ne tient pas compte du verbe d’une éventuelle complétive ou infinitive en position sujet.

Il nous reste maintenant à préciser ce que le lexique-grammaire entend par *prédictat* et *argument*. En effet, et comme cela est souvent le cas en linguistique, certains termes ont été accommodés à différentes sauces correspondant à des attaques théoriques différentes. Il en est ainsi du terme de *prédictat*. Dans une optique grammaticale traditionnelle d’ailleurs ancienne, le prédictat correspond au groupe verbal, i.e. le verbe accompagné de ses compléments, et il sert à au locuteur à assérer quelque chose à propos du groupe nominal ou sujet. Dans un sens inspiré cette fois de la logique formelle, les constituants sont de deux types : ceux servant à identifier les éléments d’un domaine, appelés variables ou termes, et ceux servant à exprimer les propriétés ou les relations entre les éléments de ce domaine, les prédictats. Dans le cas de l’exemple *Max a rencontré Lia à la plage*, les deux analyses correspondent respectivement aux deux représentations Rencontrer Lia à la plage (Max), avec *Prédicat* = Rencontrer Lia à la plage et *Argument* = Max. Et par exemple Rencontrer (Max, Lia, à la plage), avec *Prédicat* = Rencontrer, et *Arguments* = {Max, Lia, à la plage}. Cette seconde solution est celle retenue par le lexique-grammaire, comme présentant certains avantages par rapport à la première (G. Gross, 2012 : 14 sq.)¹⁴. En résumé, un schéma prédicatif est formé d’un prédictat et d’un ou plusieurs arguments, ce qu’on peut symboliser par *Pred* (*arg*₁, *arg*₂, ..., *arg*_n). Lorsqu’un texte n’est pas une phrase simple – i.e. comprend plus d’un schéma prédicatif, on parlera alors de *phrase complexe*.

Etant donnée une phrase simple, ne contenant donc qu’un seul schéma prédicatif, il peut se faire que ce prédictat se manifeste sous diverses formes morphologiques. On obtient ainsi un ensemble de phrases simples morphologiquement apparentées, différentes mais possédant le même sens. Considérons par exemple la phrase (simple) :

¹⁴ En particulier, les propriétés linguistiques générales évoquées pour qualifier les arguments sont valables aussi bien pour le sujet que pour les compléments d’objets.

- (7) Paul coupe le saucisson avec un couteau (M. Gross, 1975 : 29).

On note immédiatement qu'un certain nombre de phrases sont morphologiquement apparentées et présentent un sens proche voire identique :

- (7) (a) Paul utilise un couteau pour couper le saucisson.
(b) C'est avec un couteau que Paul coupe le saucisson.
(c) Le saucisson est coupé par Paul avec un couteau.
(d) Le saucisson, Paul le coupe avec un couteau.
(e) Le couteau avec lequel Paul coupe le saucisson.

Le passage d'une de ces phrases à l'autre est une *transformation* au sens de Harris, et représente une possibilité élargie de relier les phrases entre elles par rapport aux simples commutations de morphèmes, les seules jusque là prises en compte :

"[...] en étendant la substitution des classes des morphèmes seuls aux séquences de morphèmes, nous aboutissons à des formules qui rendent égales diverses séquences substituables l'une à l'autre dans toutes ou certaines expressions de la langue en question." (Zellig S. Harris, 1946 ; traduit in *Langages* (1968), 3, n°9 : 23-50).

En résumé, les transformations (au sens de Harris) permettent de construire des classes d'équivalence sémantiques par des procédures syntaxiques. De son point de vue, contrairement aux approches plus traditionnelles, aucun mot d'une langue n'a d'autonomie syntactico-sémantique : tout mot entre dans (au moins) une phrase simple caractéristique. De plus, toute phrase complexe est susceptible d'être analysée en termes de phrases simples. Cette vision du lexique tranche avec la position qui voit les *mots* comme des entités de nature différente des *phrases*, dans la mesure où l'optique grossienne voit en fait les mots comme renvoyant à une ou plusieurs phrases simples.

Voici, très sommairement résumée, l'approche que propose le lexique-grammaire pour l'étude systématique des entrées lexicales comme le principal défi actuel de l'étude scientifique des langues. Il s'agit de caractériser une entrée lexicale donnée par les principaux schémas prédictifs dans lesquels elle peut figurer, ainsi que les transformations au sens ci-dessus, possibles ou non. On pourra pour plus amples détails se reporter à Harris (1964, 1968a, 1968b, 1981), Boons-Guillet-Leclère (1976), Ibrahim (2003), M. Gross (1975, 1981), Balibar-Mrabti (2020), entre autres publications sur le sujet.

Quelle contribution peut apporter le lexique-grammaire au(x) problème(s) posé(s) par les différents phénomènes regroupés sous le nom de figement ?

Comme beaucoup de linguistes, M. Gross part de l'observation intuitive banale que le sens dit littéral des mots composant certaines expressions ne permet pas d'interpréter leur combinaison. On reconnaît là le fameux critère de non-compositionnalité, le critère B₅ de Bally présenté au début de cette étude, et il concerne aussi bien les locutions adverbiales (*à bride abattue*) que certains groupes nominaux (*un cul-de-sac*) ou verbaux (*mettre les pieds dans le plat*), voire des phrases entières (*Les dés sont jetés*, *Ventre affamé n'a pas d'oreille*). De plus, les recensements opérés au sein de l'équipe de M. Gross montrent que les locutions de type figé sont numériquement aussi nombreuses que les locutions libres,

et que le phénomène ne peut donc en aucun cas être considéré comme négligeable ou marginal.

La méthode proposée et mise en œuvre au sein du LADL consiste à mettre en évidence les structures figées par le biais des transformations dont nous avons parlé plus haut : certaines transformations possibles avec des structures libres ne sont pas (toujours) possibles avec les structures en apparence analogues, mais présentant des phénomènes de figement. Nous exposerons ici quelques cas à titre d'exemples, on pourra se reporter à la bibliographie pour des exposés plus complets, en particulier Danlos (1981, 1988) ; M. Gross (1982, 1988, 1993).

Nous emprunterons un premier cas à G. Gross (1996 : 12 *sq.*), et comparerons les deux phrases *L'enfant a pris le jouet* et *Luc a pris la tangente*. On note que ces deux phrases ont des structures formelles superficielles quasiment identiques, et que par ailleurs tout francophone classerait *a priori* et sans hésiter la première dans les phrases ‘normales’ et la seconde dans les phrases ‘toutes faites’. Nous montrerons en particulier que la seconde de ces phrases est à ranger dans les phrases exhibant des phénomènes de figement, et ce parce qu’elle donne lieu à des blocages des transformations (au sens de Harris) que la première en revanche autorise. Ainsi :

a) La passivation :

- (8) (a) Le jouet a été pris par l’enfant
(b) *La tangente a été prise par Luc.

b) La pronominalisation :

- (9) (a) L’enfant l’a pris (le jouet).
(b) ? Luc l’a prise (la tangente).

c) Le détachement :

- (10) (a) Le jouet, l’enfant l’a pris.
(b) La tangente, Luc l’a prise¹⁵.

d) Extraction :

- (11) (a) C’est le jouet que l’enfant a pris.
(b) ??C’est la tangente que Luc a prise.

(e) Relativation :

- (12) (a) Le jouet que l’enfant a pris.
(b) *La tangente que Luc a prise.

En première approximation donc, la distinction intuitive qu’on aurait tendance à faire entre les deux phrases – et qui tient selon nous à la présence dans la seconde d’un figement absent de la première – semble être corrélée au niveau des propriétés transformationnelles.

¹⁵ b) et c) sont diversement admis pour *Luc a pris la tangente*. SketchEngine fournit cependant l’exemple *Bon, ben je fais le jus et je la prends la tangentegrosse journée aujourd’hui*, qui a été largement accepté. On touche là, probablement, aux limites de la pertinence de ce type de critères.

D'autres propriétés confortent cette thèse, ainsi les propriétés référentielles de l'une et l'autre phrase, en particulier la non-actualisation de certains éléments. Entre autres :

a) Possibilité de détermination propre :

- (13) (a) L'enfant a pris ce jouet.
(b) Luc a pris cette tangente.

b) Possibilité d'interrogation :

- (14) (a) L'enfant a pris quel jouet ?
(b) *Luc a pris quelle tangente ?

c) Possibilité de détermination adjectivale :

- (15) (a) L'enfant a pris un jouet éducatif.
(b) *Luc a pris une rapide tangente.

On notera, à propos de l'exemple *Luc a pris la tangente*, que c'est la locution *prendre la tangente* qui est figée, et non la phrase simple tout entière. Alors que le sujet *Luc* de la phrase complète est une possibilité dans un vaste paradigme, seul le verbe *prendre* est possible pour la locution.

Un second exemple opposera *Paul a farci la dinde de marrons* et *Paul a farci son roman de citations* (Boons 1971). Il est habituel de considérer que la première phrase comporte une occurrence de *farcir* en son sens ‘habituel’, la seconde en un sens métaphorique. Comme le remarque G. Gross (1996 : 225-6), cette opération de métaphore ne se limite pas à remplacer les arguments du premier emploi par d'autres arguments, mais atteint également les possibilités transformationnelles. Certes, dans les deux exemples, les variations paradigmatisques de *farcir* semblent identiques ou à tout le moins proches : *Paul a (farcì + rempli + truffé + assaisonné) (la dinde de marrons + son roman de citations)*, mais là s'arrêtent les ressemblances. Les deux emplois divergent pour les opérations suivantes :

a) L'effacement du complément instrumental :

- (16) (a) Paul a farci la dinde.
(b) *Paul a farci son roman.

b) La paraphrase nominale avec verbe-support :

- (17) (a) Paul a mis de la farce dans la dinde.
(b) *Paul a mis de la farce dans son roman.

c) La reconstruction avec interrogation :

- (18) (a) Quelle farce a mis Paul dans la dinde ?
(b) *Quelle farce a mis Paul dans son roman ?

Ces phénomènes montrent que le glissement métaphorique n'est pas un simple déplacement sémantique, mais s'accompagne, au niveau des propriétés linguistiques, de modifications qui l'apparentent à une sorte de figement. Dans le cas présent, Boons

(1971) constate que *farce* a un statut proche de celui d'objet interne¹⁶ pour *farcir* dans le premier exemple, objet interne qu'il n'est plus dans le second exemple. Il en résulte une modification de l'interprétation sémantique : or, on sait que les processus de figement s'accompagnent toujours de modifications sémantiques. Nous en verrons d'autres cas ultérieurement. En suivant la suggestion de (Boons, 1971), on pourrait penser, sur la base de b) ci-dessus, à distinguer deux verbes *farcir* (ou deux emplois de ce même verbe) par les phrases simples qui leur sont attribuées : le *farcir* de la première phrase se verrait attribuer une phrase simple du type *No met une farce de N₁*, qui ne figureraient pas dans la description par phrases simples du *farcir* de la deuxième phrase. Dans une telle optique, certains objets internes seraient interprétables comme les traces d'une phrase simple attachée à un verbe.

4.2. La théorie *Sens-Texte* de Mel'čuk

La *Théorie Sens-Texte* (*Meaning-Text Theory*) de Mel'čuk a vu le jour à Moscou dans les années 60, principalement sous l'impulsion d'Igor Mel'čuk, de A. Žolkovskij et de J. Apresjan. Elle propose un cadre théorique de constructions de ‘modèles’ formels pour les langues naturelles. Comme les autres théories évoquées dans ce travail, elle part de l’hypothèse de base qu’on peut rendre compte des phénomènes langagiers par des modèles réglés, *i.e.* centrés sur des ensembles de règles. Notre but n'est pas ici de la résumer en quelques lignes, mais de montrer le traitement qu'elle réserve aux phénomènes de figement. Pour une présentation détaillée de cette théorie, on pourra se reporter à Mel'čuk (2015), et à Mel'čuk (2006) pour un panorama complet.

Le but de la *Théorie Sens-Texte* est l’élaboration d’une machine virtuelle permettant de traduire des *Sens* en énoncés - les *Textes*, et *vice versa*, valable *a priori* pour toutes les langues, aucune n’étant privilégiée de ce point de vue. Dans la théorie *Sens-Texte*, la modélisation d’un énoncé s’opère selon quatre niveaux de représentation, à savoir *phonologie/phonétique*, *morphologie*, *syntaxe*, et enfin *sémantique*. Ces niveaux de représentation ne sont pas isolés, et des transformations permettent de passer de l’un à l’autre. La *Théorie Sens-Texte* est donc fondamentalement une approche modulaire. Explicitement destinée à la description des phénomènes linguistiques *internes* (lexique, morphologie, collocations), sans intervention de phénomènes externes, par exemple psychologiques.

Comment fonctionne cette théorie ? La théorie *Sens-Texte* part du postulat de base qu'une langue est définie par des unités lexicales et des *fonctions lexicales* qui leur sont attachées, une *fonction lexicale* étant une représentation de la combinatoire lexicale restreinte d'une unité lexicale, *sa collocation*. Un point essentiel de la *Théorie Sens-Texte*

¹⁶ Rappelons qu'on appelle traditionnellement *objet interne* d'un verbe un constituant fonctionnel de ce verbe reproduisant en tout ou partie le sens du verbe. Il s'agit donc d'une notion à la fois syntaxique – les verbes admettant un objet interne sont souvent intransitifs, et sémantique – l'objet interne est très proche du verbe dont il reproduit le sens. Les exemples classiques d'objet interne sont des constructions comme *vivre sa vie*, *dormir son sommeil*, *chanter sa chanson*, *jouer son jeu*, *mourir d'une belle mort*, etc. Peu étudiée, cette notion reste problématique.

est qu'une langue est caractérisée par ses collocations. Enfin, les arborescences utilisées dans cette théorie pour représenter le(s) structure(s) syntaxique(s) de surface l'apparentent à une grammaire de dépendance.

Fondamentalement, la *Théorie Sens-Texte* (dorénavant *MST*) est un dispositif logique qui met en correspondance une première série d'objets formels, les *Représentations Sémantiques* (concernant le sens qu'un locuteur désire exprimer), et une seconde série d'objets formels, les *Représentations Phoniques* (concernant les textes qui d'après ce même locuteur, expriment le mieux le sens de la Représentation Sémantique correspondante). Cette correspondance entre Rsem et Rphon passe par deux intermédiaires, à savoir la *Représentation Syntaxique* (Rsynt), qui correspond aux régularités spécifiques de la *Phrase*, puis la *Représentation Morphologique* (Rmorph), qui correspond cette fois aux régularités spécifiques du *Mot*. On sait par ailleurs que le choix des lexèmes se fait selon deux axes : un axe *paradigmatique* et un axe *syntagmatique*. Pour traiter les choix paradigmatisques, la *MST* postule que les sens lexicaux sont décomposables en des sens plus simples, eux-mêmes décomposables à leur tour, jusqu'à ce que l'opération ne soit plus possible et qu'on atteigne le niveau de *primitif sémantique*, au sens de Wierzbiecka (1998), et Goddard & Peeters (2006 : 13-38).

Ces décompositions se font selon des règles présentes dans le composant sémantique de la *MST*, les *règles sémantiques lexémiques*. Les choix lexicaux, cette fois syntagmatiques, posent de redoutables problèmes, la combinatoire lexicale semblant très capricieuse. Ainsi, on opposera les deux paires *très déshabillé*/**bien déshabillé* et **très reposé/bien reposé*, ou encore *mettre en colère*/**plonger dans la colère* et **mettre en désespoir/plonger dans le désespoir* (Mel'čuk, 1997 : 23). Ces phénomènes sont connus sous le nom de *collocations*, et sont représentés dans le *MST* à l'aide de *fonctions lexicales*. Une fonction lexicale fournit, pour le sens exprimé au voisinage d'une lexie x toutes les lexies y possibles. Cette fonction lexicale sera notée $f(x) = y$. En voici un exemple, relatif à l'intensité d'un phénomène (l'écriture en gras renvoie non au mot mais au concept d'intensité) : **intensément** (dormir) = {*intensément, profondément, comme une souche, comme un loir, à poings fermés*} ; **intensément** (blessé) = {*gravement, grièvement*}.

À partir d'une représentation mentale d'une situation (la R[eprésentaion] Concept[uelle]), le locuteur sélectionne une Rsem du futur énoncé, grâce à un module RConcept → Rsem. À partir de cette représentation sélectionnée, le locuteur choisit les constructions syntaxiques et les constructions lexicales constituant la Rsynt de l'énoncé, au moyen du module Rsem → Rsynt¹⁷. C'est à ce niveau qu'intervient le traitement des phénomènes de figement, car la liberté de choix des lexies peut se voir restreinte par des facteurs linguistiques. Or, le *MST* comprend entre autres pour une langue donnée un ensemble structuré d'unités lexicales ou *lexies*, lesquelles sont assorties d'informations appropriées¹⁸. Ces lexies sont de deux types : les lexies constituées d'un seul mot pris en

¹⁷ Nous avons quelque peu simplifié le parcours, pour des raisons pédagogiques.

¹⁸ On notera une certaine parenté entre ces *informations appropriées* et les *phrases simples* de M. Gross.

un seul sens, ou *lexèmes* ; et les lexies constituées de plusieurs mots formant un syntagme, ou *phrasèmes*¹⁹. Pour être un phrasème, un syntagme doit donc être contraint, i.e. non (complètement) libre. Ces contraintes peuvent être de différentes natures linguistiques : on parlera alors de *phrasèmes sémantiques*, en réservant le nom de *phrasèmes pragmatiques* aux phrasèmes contraints par des facteurs extralinguistiques²⁰. Comme exemples de pragmatèmes, citons *Défense de stationner* ou encore *Conserver au frais* (Mel'čuk 2011, : 44-45).

Les phrasèmes sémantiques se répartissent en trois classes principales : les *collocations*, les *clichés* et les *locutions* (cf. Polguère & Mel'čuk, 2006, et Polguère, 2016 pour une application détaillée au français.)

1. Les collocations : une *collocation* est un phrasème sémantique contraint sur un seul de ses composants et sémantiquement compositionnel. Il s'agit donc d'un syntagme **AB** tel que : a) le sens de **AB** est constitué compositionnellement des sens de **A** et de **B** ; b) un des constituants, appelé *pivot*, est sélectionné, alors que l'autre est sélectionné en fonction du premier. Un exemple bien connu, emprunté à Polguère (2007), est l'opposition *soleil de plomb/chaleur accablante*. Si l'on veut en effet marquer l'intensité de la sensation dans chaque cas, le choix du pivot *soleil* fera choisir de préférence *de plomb*, et celui du pivot *chaleur* fera pencher vers *accablant*, et non l'inverse : ??*soleil accablant*/chaleur de plomb*. Notons que les verbes supports jouent souvent le rôle de **B** dans les syntagmes à verbe support comme *faire une promesse, donner une gifle, avoir du regret*, etc., où les pivots sont les substantifs, et le rôle du verbe support de les 'transformer' en verbe : *promettre, gifler, regretter*, etc.

2. Les locutions : un phrasème sémantique est une *locution* s'il est sémantiquement non compositionnel, et qu'aucun des composants n'est pivot sémantique. Par exemple *filer un mauvais coton, jeter le bébé avec l'eau du bain, se mettre le doigt dans l'œil*, etc. Le *MST* distingue trois sous-classes de locutions – les *locutions faibles*, les *semi-locutions* et les *locutions fortes*, distinction dont nous ne parlerons pas ici (sur ce point, v. Mel'čuk 2011 : 47-48).

3. Les clichés : un *cliché* est un phrasème sémantique complètement contraint et sémantiquement compositionnel. Des proverbes comme *Abondance de biens ne nuit pas* ou *Mieux vaut tard que jamais* font partie des clichés dans la mesure où ils sont totalement compositionnels – leur sens est parfaitement clair – mais sont complètement contraints : rien ne peut y être modifié, du moins dans la version officielle. Nous reviendrons sur ce problème plus avant.

Nous allons maintenant exposer brièvement la façon dont le *MST* traite les phrasèmes sémantiques, car il est clair que la notion de phrasème sémantique recoupe amplement la

¹⁹ Les phrasèmes correspondent ainsi à ce que d'autres appellent *suites polylexicales*.

²⁰ On notera que cette définition recoupe la définition générale de *syntagme figé*, à savoir syntagme polylexical (= lexie) tel que la connaissance du sens dit littéral des mots le composant ne permet pas d'interpréter leur combinaison, ou encore *phrasème sémantique*.

notion de figement telle qu'elle est largement utilisée. Cette dernière notion, que nous examinerons plus avant, repose sur une série de critères de différentes natures (la plus célèbre étant la non-passivation), ainsi que sur la notion d'*opacité sémantique*, qui ressemble à celle de non-compositionnalité, même si elle en diffère sur certains points, comme nous le verrons quand nous examinerons l'ensemble des vulgates ayant cours dans le traitement des phénomènes rassemblés sous le nom de *figement*.

Un premier point est que ce traitement des phrasèmes sémantiques, outre son intérêt théorique, est destiné à figurer dans le *DEC (Dictionnaire Explicatif et Combinatoire)*²¹ ou dans un dictionnaire de ce type, et qu'il convient à ce propos de spécifier le mode de représentation des phrasèmes dans un tel ouvrage. Deux types de représentation sont mises en œuvre selon qu'on a affaire à des locutions ou alors à des clichés et collocations.

a) Les locutions : étant d'après leur définition des phrasèmes complètement contraints et non compositionnels, ce sont des unités lexicales au même titre que n'importe quel lexème comme *dictionnaire* ou *passivation*. Il n'y a donc *a priori* aucune raison de faire figurer dans un dictionnaire une locution à l'intérieur de l'entrée d'un de ses constituants, ainsi *pomme de terre* sous *pomme* comme le font la plupart des dictionnaires usuels²². La seule différence avec un lexème est que les lexies étant polylexicales, la description d'une locution devra comporter des informations sur leur organisation syntaxique de surface (SSyntS), i.e. ce que le composant syntaxique Rsynt présente comme caractéristique de cette organisation. Par exemple les deux substantifs objets de la locution *Une hirondelle ne fait pas le printemps*, à savoir *hirondelle* et *printemps*, ne peuvent être repris par anaphore : *Une hirondelle, le printemps, elle ne le fait pas*, sauf volonté de jeu de mots. Ce type de contrainte est courant avec les locutions, mais non général. *Dans tout homme, il y a un cochon qui sommeille* supporte assez bien *Tout homme, il y a un cochon qui sommeille en lui*, dans un français, il est vrai, plus parlé qu'écrit.

b) Les collocations : étant compositionnelles, les collocations ne sont pas des lexèmes. Elles n'ont donc pas leur propre entrée dans le dictionnaire, et figurent dans une entrée qui représente leur *ancrage lexical*, dans le cas présent la base de collocation. *Soleil de plomb* figurera ainsi sous sa base *soleil*, qui est son ancrage lexical.

c) Les clichés : ils sont également compositionnels, mais leur cas est différent, dans la mesure où ils n'ont pas d'ancrage lexical unique, mais peuvent en posséder plusieurs. Ainsi le cliché cité ci-dessus *Abondance de biens ne nuit pas* est susceptible de posséder trois ancrages lexicaux, à savoir les substantifs *abondance* et *bien*, ainsi que le verbe *nuire*. Dans le cas du *MST*, ces classifications déterminent les fonctions lexicales (FL) qu'il faudra fournir pour la description d'une lexie. Ces FL figurent en effet dans le DEC, dont c'est une des grandes originalités. Pour ses auteurs, les quelque 60 FL qu'ils ont mis

²¹ Cf. Mel'čuk et alii (1984-1999). Ce dictionnaire repose sur le MST, et pose donc le problème de la représentation des unités lexicologiques choisies.

²² Assez bizarrement, le *TLF*, outre une entrée séparée pour *pomme de terre*, fait figurer ce phrasème à la fois sous *pomme* (ce n'est pas une pomme) et sous *terre* (elle pousse effectivement dans la terre).

en évidence suffisent à décrire de façon quasiment exhaustive la combinatoire des lexèmes du français ou de tout autre langue.

Voici, outrageusement simplifiée, le modèle *MST* imaginé par Mel'čuk pour rendre compte des phénomènes phraséologiques.

On aura sans doute remarqué une certaine proximité avec les travaux de M. Gross exposés *supra*. Il s'agit, dans les deux cas, de décrire de façon exhaustive les connaissances lexicales combinatoires d'un locuteur dans une langue donnée, cet objectif étant atteint par examen systématique des co-occurrences lexicales de l'unité lexicale étudiée. Dans les deux cas, ces indications figurent dans la description lexicale, que ce soit sous forme de *phrases simples* (M. Gross) ou de *fonctions lexicales* (I. Mel'čuk). Par ailleurs, les deux approches sont modulaires, i.e. prévoient des passages entre les différents composants, en particulier ceux alloués à la syntaxe et à la sémantique. Enfin, l'angle d'attaque du modèle *MST* est celui de la sémantique générative : la structure 'profonde' de départ est sémantique, et doit passer par divers modules interdépendants afin d'engendrer une structure de surface formelle.

4.3. La vulgate : les travaux autour de G. Gross

Les approches comme celles de M. Gross ou de I. Mel'čuk ont toutes deux pour caractéristique la volonté d'intégrer les phénomènes de figement (en un sens vague) dans une théorie linguistique générale. Les phénomènes phraséologiques font dans les deux cas partie intégrante des phénomènes linguistiques, au même titre que les autres phénomènes. De telles approches ne sont pas nombreuses, face à celles qui considèrent cette fois les phénomènes de figement comme des anomalies par rapport à un modèle régulier. Les déviations par rapport à une norme (celle des expressions polylexicales dites libres) doivent être mises en évidence à l'aide de batteries de tests très généralement acceptés comme pertinents. Rappelons que ces tests, qui concernent uniquement des unités polylexicales (les phrasèmes de I. Mel'čuk), sont habituellement regroupés selon trois catégories de phénomènes déjà mentionnés. Nous les résumerons comme suit :

1. L'opacité sémantique (souvent appelé *figement sémantique*) : c'est sans doute le phénomène le plus connu, et il est souvent invoqué, faute peut-être de mieux, comme définition du figement tout court: le sens global d'une expression figée n'est pas déductible du sens des éléments qui la composent formellement. Sont par exemple sémantiquement opaques des 'locutions' comme *mettre les pieds dans le plat, se mettre le doigt dans l'œil, à la Saint-Glinglin, à la va-comme-je-te-pousse, manger les pissenlits par la racine, de fil en aiguille, marcher sur des œufs*, etc., sans oublier l'incontournable *casser sa pipe*. On aura reconnu là le fameux *Principe de non-compositionnalité*, dont nous verrons plus avant les limites.
2. Le blocage de certaines manipulations (parfois appelé *figement transformationnel*). Le plus célèbre de ces critères, et l'un des plus anciens – nous verrons plus loin sa fiabilité – est la passivation. Ainsi, hors la volonté de jeux de mots, *Le cheval a pris le mors aux dents* n'admet pas la forme passive *Le mors aux dents a été pris par le cheval*, pas plus

que *Max n'a pas inventé le fil à couper le beurre* n'accepte *Le fil à couper le beurre n'a pas été inventé par Max*. Mais d'autres blocages ont été fréquemment convoqués, ainsi ceux mentionnés plus haut (G. Gross, 1996) : la reprise pronominale, le détachement, l'extraction, et la relativation, pour ne citer que les plus courants.

3. La non-actualisation des éléments (parfois appelé *figement référentiel*) : ce critère, rappelons-le, concerne essentiellement le groupe nominal et les possibles variations paradigmatisques de ses éléments, ainsi que les potentialités référentielles. Ainsi, dans la locution *passer comme une lettre à la poste*, l'article indéfini *une* ne fait en aucune façon partie du paradigme habituel des déterminants, d'où le contraste *Ma demande est passée comme une lettre à la poste / Nos demandes sont passées comme (une lettre + ??des lettres) à la poste*²³. On pourrait penser au cas de *mettre la clé sous la porte / mettre les clés sous la porte*. En fait, l'impossibilité d'accord avec le sujet montre qu'il s'agit de deux variantes et non d'une possible variabilité d'une seule locution : *En faillite, cette entreprise a mis les clés sous la porte / En faillite, ces entreprises ont toutes mis la clé sous la porte*.

L'ensemble de ces critères a fait l'objet de nombreuses études, qu'il serait vain de vouloir citer de façon exhaustive. En nous bornant à quelques publications à visée générale, mentionnons, outre les auteurs déjà cités, G. Gross (1996), Mejri (1997), Grossman-Tutin (2003), Svensson (2004), pour le français ; Palma (2007), pour une étude contrastive français-espagnol ; Ruiz-Gurillo (2001), Montoro (2006), Mogorrón Huerta (2008) et surtout García Page (2008), pour l'espagnol, et enfin Quiroga (2006) pour une étude contrastive espagnol-italien.

On aura par ailleurs remarqué que les critères évoqués dans ces travaux et largement acceptés sont loin d'être originaux, comme nous l'avons déjà noté *supra* :

- a) Le critère de non-compositionnalité recoupe les critères B₅ (l'oubli du sens des éléments) et B₆ (la présence d'un archaïsme)²⁴ de Bally.
- b) Le blocage de certaines manipulations (figement transformationnel) correspond au critère B₂ (ces mots sont disposés dans un ordre invariable, et ne peuvent être séparés par d'autres mots) ;
- c) Enfin, la non-actualisation des éléments (figement référentiel) correspond peu ou prou au critère B₃ (aucun des mots du groupe ne peut être remplacé par d'autres mots).

²³ La répartition des déterminants dans ce type d'expressions semble en partie erratique, et n'a jamais été étudié. Ainsi : *écrire comme un pied mais penser avec les pieds* ; *vif comme l'éclair mais lent comme une tortue* ; *plat comme une limande mais gros comme le bras*, etc.

²⁴ L'hypothèse de l'archaïsme a fréquemment été évoquée pour se sortir d'affaire. Ainsi, les constructions du type verbe support + nom prédictif comme *donner ordre*, étudiées dans Giry-Schneider (1978), ont été souvent présentées comme un archaïsme issu de l'ancien français dans l'ignorance totale des données diachroniques, mais également synchroniques, comme exposé dans Anscombe (1986).

Signalons également que l'hypothèse de la polylexicalité n'est autre que B₁ (le groupe est composé de plusieurs mots séparés par l'écriture), et que B₄ (l'équivalence de l'unité examinée avec un mot simple) est un des principes implicites de la plupart des dictionnaires, qui expliquent un idiome polylexical comme *ne pas souffler mot* par un lexème monolexical 'se taire'.

Si donc on tente de présenter un bilan provisoire des travaux examinés ci-dessus et qui comprennent la ligne majoritaire des recherches sur les phénomènes de figement ou d'idiomaticité, on peut dire que deux traits principaux les caractérisent :

- a) Les phénomènes dits de figement sont considérés comme hors système. Ils ne peuvent être décrits à l'aide des seules règles du système choisi, en particulier pour ce qui est de la combinatoire sémantique, et ne relèvent donc pas de la morphologie usuelle.
- b) Ces phénomènes de figement constituent une catégorie homogène, en ce sens qu'ils sont identifiables au travers de critères spécifiques, ceux précisément que nous venons d'exposer.

Pour ce qui est du traitement de tels phénomènes dans le cadre d'une théorie linguistique, la tendance majoritaire consiste à les étudier à part du système général choisi et donc de sa représentation lexicale, au travers de descriptions spécifiques. Certaines théories plus minoritaires proposent des modèles qui intègrent les unités phraséologiques dans le lexique, en leur adjoignant une 'carte d'identité' qui spécifie les contraintes de divers types auxquelles l'unité donne lieu, sous forme de *phrases simples* (M. Gross) ou de *fonctions lexicales* (Mel'čuk).

4.4. Quelques concepts additionnels : semi-figement, degré de figement, défigement

Les tenants de la vulgate telle qu'elle vient d'être exposée se sont vus très vite confrontés à un problème redoutable qui les a conduits à moduler certaines hypothèses. Le fait le plus immédiatement visible est qu'il existe des constructions apparemment non-compositionnelles (donc méritant l'appellation de 'figées'), de formes similaires, et qui cependant ne satisfont pas les mêmes critères de figement. Considérons par exemple les deux locutions *mettre les pieds dans le plat* et *mettre la clé sous la porte*. Elles sont clairement non-compositionnelles, et diffèrent pourtant quant aux possibilités d'actualisation de l'article : *mettre les pieds dans le plat* / **mettre le pied dans le plat* mais *mettre la clé sous la porte* / *mettre les clés sous la porte* ainsi (SketchEngine, 17/01/2023) :

D'après les professionnels de l'hôtellerie, un tiers des hôtels encourt le risque de mettre les clés sous la porte...

Ce n'est pas le seul cas. Ainsi, *jeter le bébé avec l'eau du bain* se passive sans trop de problèmes : *Le bébé avait été jeté avec l'eau du bain* (SketchEngine, 30/01/2025) :

On se rappelle que la thématique Santé Environnement avait été tout simplement écartée de la dernière Conférence Environnementale. **Le bébé avait été jeté avec l'eau du bain** et le Rassemblement Pour la Planète dégagé sans façon pour éviter de faire trop de peine à certains lobbies....

alors que la même transformation semble plus difficile avec *jeter le manche après la cognée*, de forme pourtant proche. D'où l'idée que le figement ne peut être réduit à des considérations purement distributionnelles, et que certaines relations entre éléments peuvent affecter (ou non) la combinabilité de ces éléments dans une structure donnée. C'est ainsi qu'on a parlé de *semi-figement* García-Page (1999), (Balibar-Mrabti, 1994, 2005 ; Lamiroy & Klein, 2005 ; Mejri, 2005) et de *degré de figement* (G. Gross, 1996 : 16 sq. ; Schapira, 1999 : 9 sq.) et même de *continuum de figement* (G. Gross, 1996 : 17 ; Mejri, 1997 : 36 sq.). Il s'agit de caractériser le phénomène qui fait que les expressions considérées comme figées peuvent posséder certains 'degrés de liberté' difficiles à caractériser. La position de Schapira est parfaitement claire sur ce point :

« ... À côté de ces critères distinctifs, qui se vérifient pour toutes les expressions figées, il en existe d'autres, moins rigoureux, qui s'appliquent selon le degré de figement plus ou moins élevé de la séquence. En effet, le figement est un **phénomène scalaire**, ses produits pouvant présenter une rigidité plus ou moins intense. Il est intéressant de noter que la rigidité du figement va croissant des unités lexicales aux énoncés phrasiques, qui, eux, accusent un figement complet : le proverbe, le dicton, le slogan, etc., ne tolèrent aucun changement formel. Les locutions, en revanche, n'obéissent pas toujours aux critères de figement figurant ci-dessus ... ».

En fait, ces thèses reposent sur deux thèses sous-jacentes que nous examinerons dans un paragraphe ultérieur, et qui sont les suivantes : a) Les expressions figées sont une catégorie homogène par rapport à des critères distinctifs que tous ses membres vérifient ; b) Le degré de figement augmente avec la longueur de la séquence figée considérée. Nous montrerons ultérieurement le caractère erroné de ces deux thèses. Un premier point est que les notions de semi-figement et de degré de figement sont peu claires : nous n'avons pu trouver une quelconque définition qui serait du type « la locution *X* est plus figée que la locution *Y* si les propriétés P_1, P_2, \dots , sont satisfaites ». On nous permettra par ailleurs de critiquer l'usage abusif qui est fait du terme de *continuum*, qui ne peut s'appliquer à des termes discrets.

Un dernier concept a été évoqué pour mettre en évidence un éventuel figement : le *défigement*. Il s'agit, selon G. Gross (2012 : 38) :

« ... de transgressions opérées sur des suites générées par des règles, mais comme une activité ludique ... ».

Par exemple *Qui trop embrase, mal éteint*, sur le modèle de *Qui trop embrasse, mal étreint*, exemple de Barta (2005) cité par Palma (2007 : 161)²⁵. On trouvera d'ailleurs chez Palma (2007 : 155 sq.) un panorama complet de la question, bien que l'auteur regrette que :

« ... il n'existe pas d'étude complète visant à dégager clairement les procédures à la base de ce jeu expressif ... ».

²⁵ Ce procédé n'est pas nouveau, et avait déjà été largement exploité par Balzac, auteur entre autres de *Les petits poissons font les grandes rivières*. Sur ce sujet, on pourra lire Navarro Domínguez (2000).

L'idée, que l'on trouve dès Zuluaga (1980), est que la possibilité de détournement peut être utilisée comme critère de figement, d'où le nom de *défigement*. Alors que la variation s'appuie sur le figement et conserve le sens de départ – ainsi *Qui vit par le fer, périra par le fer/Qui vit par l'épée, périra par l'épée* – le défigement altère la signification de la combinaison de départ. Ce critère est évoqué chez plusieurs auteurs, en particulier, outre ceux déjà cités, Mejri (1997b ; 2011 : 74-76), Zuluaga (1999)²⁶, Schapira (1999 : 155 : sq.), Svensson (2004), García Page (2008), G. Gross (2012 : 204), Ben Hamor Ben Hamida (2015), parmi beaucoup d'autres. Bien que Svensson (*op. cit.* : 12 *sq.*) consacre tout un paragraphe au défigement, elle en critique la pertinence comme critère de figement. Elle remarque en effet, à juste titre, que pour pouvoir parler de défigement, il faut avoir auparavant constaté de façon indépendante le figement de la combinaison examinée. Sinon, il y aurait circularité. Elle note par ailleurs à ce propos le caractère flou de la dénomination de *critère*, observation sur laquelle nous reviendrons.

4.5. Constat critique et conclusion(s)

Nous voudrions dans ce paragraphe dresser un premier bilan de l'état des recherches sur les problèmes de figement, et qui correspond à la situation théorique générale jusqu'à l'aube des années 2000, époque à laquelle apparaîtront ce qu'on a appelé *les grammaires de construction* et démarches apparentées, alternative aux théories classiques du figement dont nous parlerons plus loin. Nous ne nous proposons pas de procéder à une étude détaillée des différents problèmes que pose la notion de figement – une telle étude est faite dans le très complet ouvrage de Svensson (2004), ainsi que dans García-Page (2008) pour l'espagnol. Plus modestement, nous voudrions mettre le doigt sur un certain nombre de faiblesses théoriques caractéristiques de la 'thèse classique' du figement. À commencer par la définition de la notion elle-même de *figement*.

4.5.1. Figement et non-compositionnalité

La problématique du figement part de l'observation apparemment indiscutable que le sens de certaines combinaisons lexicales – *les idiomes*, n'est pas calculable à partir du sens des mots qui les composent. Le phénomène est transcatégoriel, et il peut s'agir de locutions adverbiales (*à tort et à travers*), de locutions verbales (*faire le pied de grue*), de noms composés (*pince-monseigneur*), de locutions prépositionnelles (*à toutes fins utiles*), voire de phrases entières (*La mariée était trop belle*). C'est ainsi que les proverbes sont généralement considérés comme l'exemple type de phrases figées (Norrick, 1985 : 31 ; Mejri, 1997a : 242 ; Schapira 1999 : 9 ; Conenna, 2000 : 29 ; Kleiber, 2000 : 40 ; Svensson, 2004 : 23, etc.). Comme nous le verrons plus loin, cette thèse est contestée par d'autres, ainsi Palma (2004 : 122), Anscombe (2011 : 36 *sq.* ; 2012 : 30 *sq.*). Considérons par exemple le proverbe bien connu *Qui n'entend qu'une cloche n'entend qu'un son*. Il possède un *sens compositionnel*, à savoir 'si on sonne une seule cloche, on n'entendra nécessairement qu'un seul son', qui frôle la lapalissade. En revanche, le *sens figuré* de cette même tournure, dans la terminologie de Tamba-Mecz (2014b), est que 'si on se borne à une seule opinion, on ne peut se faire une idée exacte de la situation

²⁶ Qui utilise le terme espagnol *desautomatización* pour désigner le défigement.

concernée'. On constate sur cet exemple simple qu'effectivement, la relation entre les deux sens compositionnel et figuré ne va pas de soi, et pose la redoutable question de sa nature. Si notre proverbe est considéré comme figé, et qu'il a par ailleurs un sens compositionnel et un sens figuré non-compositionnel, qu'est-ce qui est figé dans cette disposition ? Si nous admettons que le sens figuré d'une expression dite figée découle du sens compositionnel de base par le biais d'une convention donnée, la solution la plus raisonnable est de considérer que c'est cette convention qui est en quelque sorte figée. Deux remarques à ce propos : la première est que les conventions qui font passer d'une valeur sémantique à une autre ne sont pas nécessairement figées en ce sens. Ainsi, les linguistes qui ont étudié les actes de langage ont proposé de rendre compte du fait que *Il fait beau* en arrive à signifier dans certains contextes *Allons nous promener*. Pour ce faire, ils ont postulé l'existence d'une convention du type 'Quand il fait beau, on va se promener', assortie d'une *Loi de discours* affirmant qu'énoncer A revient à signifier à l'interlocuteur de chercher la phrase B visée par le locuteur par le biais d'une convention, et acceptable au vu des conditions d'énonciation. Notons cependant la différence d'avec nos expressions figées. D'une part, dans notre exemple du beau temps de promenade, il y a un savoir commun qui joue le rôle de convention. D'autre part, l'application de la loi de discours est facultative : on peut constater le beau temps sans nécessairement vouloir signifier le désir de promenade. Or dans le cas de nos expressions figées, le sens figé est le seul possible. Même si le sens compositionnel est toujours présent, il ne peut intervenir que sous l'angle d'un jeu de langage. Il n'est par ailleurs pas certain que toute expression figée repose sur un savoir commun restituables au travers de règles linguistiques, ainsi *être bête comme ses pieds*, expression dans laquelle la relation de la bêtise aux pieds est quelque peu obscure. Le segment *comme ses pieds* a, comme très généralement dans ce type de comparatives dites 'à parangon', une simple valeur d'intensification résultant d'un mécanisme difficilement identifiable et dont l'origine se perd souvent dans la nuit des temps. Il est même loisible de se demander s'il existe réellement un sens compositionnel dans certains cas extrêmes, ainsi *avoir maille à partir* (*avec quelqu'un*), où *maille* désignait à l'origine une monnaie valant un demi-denier, i.e. la plus petite valeur possible, d'où l'impossibilité de *la partir* 'la partager'. Or cette expression n'a pas de sens compositionnel, les deux mots *maille* et *partir* n'existant plus avec les sens mentionnés en français contemporain. Nous avons ainsi l'exemple d'une locution verbale qui fait bloc, qui satisfait tous les critères de figement que nous avons mentionnés, mais qui, n'ayant pas de sens compositionnel, n'est pas figée, ou en tout cas, pas dans le même sens que d'autres. On voit apparaître le problème général de la définition habituelle du figement : elle repose à la fois sur un divorce entre un sens compositionnel et un sens non-compositionnel, et sur un certain nombre de critères censés aller dans le même sens. Or dans des cas comme *avoir maille à partir* et bien d'autres, il n'y a pas à proprement parler de sens compositionnel, pour la simple raison que le sens non-compositionnel est le résultat d'une évolution en diachronie qui a fait disparaître ce possible sens compositionnel— donc pas de figement selon ce critère, alors que, selon d'autres critères comme la passivation ou la reprise pronominale, on ne peut que conclure à un figement. D'où contradiction, liée à cette non-prise en compte de la dimension diachronique du phénomène. Signalons enfin que le figement tel qu'il est défini fait intervenir la notion de polylexicalité et donc de mot, ce qui nous renvoie à un problème redoutable, celui de la

définition des unités lexicales. Si l'on admet qu'une théorie linguistique explique les règles qui régissent la combinaison de 'briques' élémentaires pour former des unités plus complexes, il reste à définir ces briques élémentaires que sont les mots. On se heurte alors à la difficulté bien connue de la définition de ce qu'est un mot, et en premier lieu à une représentation orthographique qui ne facilite pas les choses. Ainsi, le *TLF* écrit *pince-monseigneur* et lui donne une entrée indépendante, alors qu'il écrit *moulin à vent*, combinaison qui figure sous *moulin*. Pourtant, une pince-monseigneur est un type de pince, ce qui fait qu'il s'agit d'une suite polylexicale au même titre que *moulin à vent*. Outre ce problème – le lexique est un construit et non un observable, un second problème est celui du sens des mots. En effet, le choix d'un sens va de pair avec des problèmes de polysémie. Prenons par exemple les deux expressions *s'asseoir à la table* et *s'asseoir à table*, qui n'ont pas le même sens : entre autres, dans la seconde, *table* signifie exclusivement 'meuble sur lequel s'effectuent les repas', spécification qu'on retrouve dans l'opposition *être à la table (de)* et *être à table* 'être en train de manger' (sans compter le sens argotique *se mettre à table* 'avouer'). Si, à l'instar de beaucoup de dictionnaires, on prend comme sens de base le sens concret de 'meuble', alors le sens 'abstrait' de 'repas' doit être considéré ou bien comme une polysémie, ou bien comme réductible par le biais d'une métonymie. Si on vise en revanche une unique valeur sémantique de base, on doit alors faire face à la difficulté de préciser cette valeur²⁷, ce qui implique de passer par des structures non explicites en surface. On voit ainsi apparaître, à l'occasion de la notion de figement, les nombreux problèmes que pose l'adoption pure et simple de notions issues non pas d'une réflexion linguistique, mais d'une tradition grammaticale qui tend à se restreindre à la structure de surface, et dont la démarche ne passait pas toujours par l'élaboration d'une théorie scientifique (cf. M. Gross, 1976).

4.5.2. Les critères de figement

Les nombreux linguistes ayant abordé les problèmes regroupés sous la rubrique *figement* ont sans doute eu très tôt le sentiment que le critère de non-compositionnalité était loin d'être satisfaisant, et devait être au mieux considéré comme un indice de figement plutôt que comme une preuve absolue. Il y a certes des cas où ce critère s'applique parfaitement : ainsi la locution prépositive *avec les doigts de pied en éventail* a un sens compositionnel, certes bizarre, mais effectif. Son sens figuré, à savoir 'dans un état de totale relaxation' est difficilement déductible des sens habituellement attribués aux composants lexicaux de l'expression. Une remarque identique vaut pour (*faire quelque chose*) *les doigts dans le nez* ou encore *effeuiller la marguerite, retirer une (sacrée) épine du pied, pendant cent sept ans, mettre les petits plats dans les grands*, etc. Mais il y a par ailleurs des cas abondants où le critère de non-compositionnalité ne s'applique pas ou mal. Un premier cas, que nous avons mentionné, est celui d'unités apparemment

²⁷ Cette valeur spécifique est due à la présence d'un article zéro, qui requiert la possibilité d'une lecture de type processif du substantif auquel il est accolé (cf. sur ce point Anscombe 1991). Autres exemples analogues : *au pied/à pied, au terme/à terme, au titre/à titre*, etc.

polylexicales comme *avoir maille à partir*²⁸, qui ne possèdent pas de sens compositionnel : *n'avoir ni sou ni maille, ne pas être dans son assiette, tailler une bavette, être tiré à quatre épingle*s, *avoir le coup de foudre, chercher noise, faire la pige, à la Saint-Glinglin, des espèces sonnantes et trébuchantes, de but en blanc*, etc. À l'opposé, le français, utilise un grand nombre de 'locutions', qui paraissent fixes (et donc figées), pour lesquelles l'opposition sens compositionnel/sens figuré semble plus floue. Ainsi : *sain et sauf, faire la sourde oreille, tenir table ouverte, être à tu et à toi, être dans le plus simple appareil, avec armes et bagages, rire dans sa barbe, trouver porte close, à travers champs, être aux commandes, corps et biens, faire une crasse, jurer ses grands dieux, n'avoir ni foi ni loi, dormir à poings fermés*. C'est sans doute parce que la limite entre les deux catégories n'est pas toujours nette que les critères mentionnés plusieurs fois antérieurement ont été imaginés. Avant d'examiner le fonctionnement de ces critères, nous voudrions revenir plus en détail sur le problème mentionné *supra* de la notion de mot, plus exactement de celui de la sémantique du mot. Nous l'illustrerons par le cas des composés N_1 à N_2 , étudiés dans Anscombe (1990, 1999), en reprenant l'analyse de Anscombe (2019a). Comme déjà noté dans ce dernier texte, l'étude détaillée de ce phénomène particulier met en relief une caractéristique importante de la notion de figement : elle ne peut avoir de sens qu'à l'intérieur d'une théorie qui confère un sens bien précis à la notion de mot. C'est d'ailleurs le choix d'une théorie bien particulière qui conduit G. Gross (1988, 1996) à qualifier de plus ou moins figés certains de ces composés. Pour le montrer, nous comparerons de ce point de vue les trois composés *moulin à vent, moulin à huile* et *moulin à café*. Si dans une théorie linguistique, nous nous bornons aux définitions des mots présentées dans par exemple le *TLF*, à savoir *moulin* = 'machine à broyer, à piler, à pulvériser diverses substances, et à extraire certains produits' ; *vent* = 'déplacement d'air plus ou moins important ressenti à la surface du globe' ; *huile* = 'substance grasse, onctueuse et inflammable, d'origine végétale, animale ou minérale, généralement liquide au-dessus de 10° centigrade, se figeant ou se solidifiant à une température inférieure, employée dans différents usages' ; enfin *café* = 'graine du cafier', alors les conclusions s'imposent. *Moulin à vent* n'est pas compositionnel, non plus que *moulin à huile*, et seul *moulin à café* est compositionnel. Pour que par exemple *moulin à huile* soit compositionnel, il faudrait que le mot *huile* comporte l'indication que certaines huiles sont obtenues à partir par exemple des olives. Le phénomène s'éclaire lorsqu'à l'instar d'Anscombe (1990, 1999), on examine les anaphores possibles sur ces composés. Or, non seulement elles sont possibles sous certaines conditions, mais les anaphores associatives étudiées dans (Fradin, 1984) et appliquées à ces trois composés (Anscombe, 1990) montrent que dans chaque cas, N_1 et N_2 sont reliés par un ou plusieurs prédicats non visibles en 'structure de surface', et qui spécifient le type de relation (locative, stative ou processive) qui existe entre les deux membres. De ce point de vue, et sans entrer dans le détail d'une analyse qui n'a pas sa place ici, les trois composés sont compositionnels, à condition de considérer que le lexique intègre les relations qui existent entre certains mots, et qu'on peut représenter par des phrases stéréotypiques. Les mots comme *moulin* renvoient à des savoirs communs, des *stéréotypes*, représentables en

²⁸ Nous disons 'apparemment', car dans ce cas précis, la polylexicalité est purement orthographique. On ne changerait rien au phénomène en écrivant *maillapartir*, de la même façon qu'on écrit *gentilhomme* et *lieudit*.

particulier par des phrases de type générique (Anscombe, 2001, 2011 : 22-27). Dans le cas de *moulin*, par exemple, sa description lexicale comporterait quelque chose comme *Certains moulins fonctionnent avec le vent*, ce qui permettrait et expliquerait *moulin à vent*. Si *huile* est décrit comme comportant le stéréotype *On fait de l'huile avec certaines olives*, on obtient *moulin à huile* à partir de *moulin* 'machine à broyer ...', etc. On voit ainsi que la propriété de compositionnalité ou aussi bien de non-compositionnalité n'est pas un observable, mais s'inscrit dans et dépend donc de la théorie lexicale choisie comme cadre de l'étude. Le lexique d'une langue n'est pas extérieur à la théorie linguistique choisie, mais en fait partie intégrante, et résulte par là-même d'une série d'hypothèses. Et le cas particulier des noms composés montre que les relations entre phraséologie et morphologie sont loin d'avoir été totalement explorées, problème exposé en détail dans le cas de l'espagnol par García-Page (2015). Enfin, la notion de *mot* doit être spécifiée dans un cadre faisant intervenir une dimension syntaxique et sémantique comme celle évoquée dans ce paragraphe, et qui n'est pas sans rappeler les *phrases simples* de M. Gross.

Si donc on ne peut que se féliciter que le concept quelque peu flou de non-compositionnalité ait été remplacé par la série de critères exposés antérieurement, nous ne sommes pas pour autant sortis de l'auberge. Deux problèmes apparaissent à ce propos : d'une part, la notion de critère elle-même. Que signifie, du moins en linguistique du figement, le fait que satisfaire pour une forme donnée F à la condition A est un critère pour F d'appartenance à une catégorie B ? Si cette formulation correspond effectivement à la pratique courante, elle est en fait ambiguë entre une condition nécessaire et une condition suffisante. Il ne peut s'agir d'une condition nécessaire : ainsi, bien des suites polylexicales sont considérées comme figées (leur sens figuré est non-compositionnel) et sont cependant passivables. Ainsi *mettre les petits plats dans les grands* (*SketchEngine*, 17/02/2025) :

« ... Cette année encore, *les petits plats ont été mis dans les grands* avec la présence de Justin Bieber. La star canadienne, à peine 17 ans et déjà empêtrée dans une sombre histoire de paternité, est l'invité d'honneur de cette édition 2011... ».

Ou encore *vendre la mèche* (*SketchEngine*, 17/02/2025) :

« ... Aucune annonce ministérielle officielle n'a été faite sur les télévisions et les radios : *la mèche a été vendue* lors de l'université d'été de la Défense, dont la 14e édition s'est tenue à Palaiseau, en région parisienne, les 5 et 6 septembre ... ».

Ce n'est pas non plus une condition suffisante : certaines locutions qui ne se passivent pas n'en sont pas pour autant figées, comme *courir un grand danger* : *Max court un grand danger/*Un grand danger est couru par Max*, phénomène qui tient à l'aspect statif de *courir* dans cette locution²⁹. On voit où gît la difficulté : à partir d'une série d'intuitions, on tente de délimiter une catégorie linguistique, celle des expressions figées. Le critère de la non-compositionnalité du sens figuré étant quelque peu flou, on l'appuie par divers critères syntaxiques, sémantiques et référentiels. Dans cette optique, si une

²⁹ Il est bien connu que les énoncés statifs se passivent mal. On comparera de ce point de vue *Du balcon, on a vu l'éclipse* et *Du balcon, on voit la mer* : *Du balcon, l'éclipse a été vue/*Du balcon, la mer est vue*.

expression satisfait un nombre suffisant de ces critères, elle sera considérée comme 'raisonnablement' figée, sans qu'aucune de ces propriétés ne soit ni nécessaire, ni suffisante. Il s'agit donc d'indices, et non de preuves *stricto sensu*.

Le second problème est que l'hypothèse d'une catégorie repérable qui serait celle des expressions figées se heurte à (au moins) deux obstacles. Le premier, que nous avons signalé à plusieurs reprises, est que deux expressions de même forme et toutes deux réputées figées, peuvent ne pas satisfaire les mêmes critères de figement. Un cas extrême est celui de la paire *jeter les hauts cris/pousser les hauts cris*, toutes deux signifiant 'protester avec véhémence' et considérées comme figées, car de sens figuré non-compositionnel, bien que ce sens figuré soit assez transparent. Or, seule *pousser les hauts cris* admet la passivation, ainsi (*SketchEngine*, 03/02/2025):

La meilleure démonstration de l'efficacité du paquet neutre en matière de lutte contre le tabagisme, *ce sont les hauts cris poussés par l'industrie du tabac qui la fournissent !*

Dans de tels exemples, *jeter* n'est pas accepté. La raison semble en être d'ordre diachronique : les étapes antérieures du français ont dit successivement :

crier les hauts cris (*Dict. Acad.*, 1694, 1718, 1740, 1762) ;
faire les hauts cris (*Mémoires de Saint-Simon*, Hachette, T. 1, 1856, p. 74) ;
jeter les hauts cris (*Dict. Acad.*, 1740, 1762, 1798, 1835, 1878, 1932-35) ; et enfin
pousser les hauts cris (*Dict. Acad.*, 1835, 1878, 1932-35).

Seuls *jeter les hauts cris* et *pousser les hauts cris* ont subsisté en français contemporain, le premier étant jugé culte voire archaïque par les francophones interrogés. *Pousser les hauts cris* est plus fréquent que *jeter les hauts cris*, surtout à l'écrit, d'après les corpus que nous avons consultés. On voit ainsi réapparaître ce facteur très largement négligé dans les études phraséologiques : la dimension diachronique, que nous avons évoquée plus haut. L'exemple traité n'est d'ailleurs pas un cas isolé : alors que *rompre la glace* et *casser sa pipe* ont des formes très proches et ont toutes deux un sens figuré non-compositionnel – seule la première locution admet certaines formes de passivation :

mais à la deuxième réunion de médiation, l'employeur l'a, à nouveau, appelée par son prénom et *la glace a été rompue ...* » (*SketchEngine*, 20/02/2025)

et non la seconde, sans raison apparente. Le second obstacle est que dans plusieurs cas de figement avéré, les critères de figement ne semblent pas corrélés. Il en est ainsi pour les comparatives dites *proverbiales* ou encore à *parangon* : *être pauvre comme Job*, *être mince comme une feuille de papier à cigarette*, *être fin comme du gros sel*, *être fier comme Artaban*, *être raide comme la justice*³⁰, etc. Très largement non-compositionnelles, elles sont quasi-unaniment réputées figées, ainsi M. Gross (1984 : 238), G. Gross (1996 : 119), Mejri (1997a : 444 *sq.*), Lépinette (1999), et enfin Schapira (1999 : 3), pour qui elles sont figées car :

³⁰ Elles ont fait l'objet de nombreuses études. Cf. en particulier Leroy (2007), Anscombe (2011, 2020) pour le français ; Pamies (2005), Garcia-Page (2007, 2011) pour l'espagnol ; Lépinette (1999), Anscombe (2020) pour une étude contrastive français-espagnol, Vietri (1990) pour l'italien.

« [...] comme nous l'avons vu, une expression stéréotypée se définit en premier lieu comme une expression figée [...] ».

Cette opinion n'est pas partagée par d'autres linguistes, ainsi Leroy (2007) et Anscombe (2008), qui contestent dans ce cas précis le lien entre figement et idiomatique³¹. En effet, ces structures sont productives (et beaucoup d'écrivains ne se sont d'ailleurs pas fait faute d'exploiter cette possibilité), comme le montrent les exemples suivants, repris de Anscombe (2011 : 31 sq.) : *pâle comme (un linge + la mort + un mort + une feuille de papier + une endive)*, *beau comme (la nuit + Apollon + Adonis + un petit Jésus + le jour + un dieu + un camion + tout)*, *raide comme (la justice + un piquet + un manche à balai + un barreau de chaise + des baguettes de tambour + un passe-lacet)*; (*sale + fier + moche + vexé) comme un pou*, (*sale + saoul + bourré + bête) comme un cochon*, (*fort + gras + gros + épais) comme un bœuf*, etc. Or, la productivité va à l'encontre des critères de figement paradigmique et éventuellement syntagmatique.

Une première conclusion s'impose donc : il ne semble pas y avoir de relation systématique entre le caractère figuré d'une expression et les critères de figement habituellement considérés, lesquels fournissent des résultats divergents – il n'y a pas de correspondance directe, et par ailleurs hétérogènes – les critères peuvent être contradictoires.

4.5.3. Problèmes de corpus

Un problème qui n'est que très rarement évoqué dans les études sur le figement et plus généralement en linguistique, est celui de la validité des conclusions tirées de l'observation de corpus de données. En effet, ces conclusions reposent sur trois hypothèses la plupart du temps implicites, et régulièrement ignorées par les utilisateurs de 'grands corpus outillés', comme si l'automatisation de la tâche effaçait l'interprétation humaine des résultats qu'elle procure. S'autoriser en effet à fonder des conclusions théoriques sur un corpus *C* revient à faire trois hypothèses :

- a) *C* est exhaustif, i.e. l'ensemble des manifestations du phénomène étudié est entièrement illustré dans *C* ;
- b) *C* est pertinent, i.e. les manifestations du phénomène étudié présentes dans *C* montrent les relations du phénomène étudié avec d'autres phénomènes et leurs propriétés ;
- c) *C* est homogène, i.e. les manifestations observées du phénomène étudié relèvent d'un ensemble unique de règles. C'est ce dernier point que nous voudrions commenter, en nous limitant à l'exemple du français.

En effet, la plupart des études sur le figement montrent les caractéristiques suivantes :

- a) L'établissement de liste de locutions et autres phénomènes phraséologiques a majoritairement recours aux dictionnaires et autres recueils disponibles, la plupart du temps sans recul critique. Or, pour tout ce qui est phraséologique, les dictionnaires ont la

³¹ Pour les problèmes analogues de la comparative à parangon en espagnol, cf. Pamies (2005) et García-Page (2011).

fâcheuse manie de procéder par compilation de dictionnaires et d'ouvrages antérieurs. D'où la fréquente présence dans ces mêmes dictionnaires et recueils d'expressions archaïsantes ou même carrément obsolètes, sans compter la transmission quasi-religieuse d'erreurs de tous types. Ainsi, le recueil de locutions de Sommant (2003), à visée pourtant moderne, nous présente toute une série de locutions vieillies voire obsolètes comme *repasser les plats*, *avoir du yaourt dans la tête, pour le quart d'heure*, *y avoir un cheveu*, *s'en falloir d'un doigt*, *huile de bras*, *avoir une mine de chat fâché*, *s'en mettre plein la lampe*, etc., sans fournir le moindre corpus de référence.

b) Les travaux utilisant plusieurs sources d'exemples ne se posent jamais la question de l'homogénéité globale de ces corpus. Or, nos langues possèdent la particularité d'être à tradition écrite, y compris pour éventuellement représenter de l'oral. Entre les divers registres de l'écrit et les divers registres de l'oral représentés, les variations peuvent être importantes, voire antagonistes. Considérons par exemple le très connu proverbe *Tout vient à point à qui sait attendre*, version officielle contemporaine de *Tout vient à point qui sait attendre* – laquelle remonte à une époque où le *qui* pouvait avoir valeur de *si on* – valeur qu'on rencontre encore mais qui semble en voie de disparition. Tous les francophones interrogés connaissent la version 'officielle', les plus âgés mentionnent également la version 'archaïque' *Tout vient à point qui sait attendre*. Les dictionnaires usuels suivent cette ligne, sans d'ailleurs la justifier : *Tout vient à point à qui sait attendre* (Gd Robert, Larousse, TLF) ; *Tout vient à point qui sait attendre* (Gd Robert, Larousse). Les choses se gâtent au niveau des recueils, qui proposent un plus grand éventail de variantes : *Tout vient à point à qui sait attendre* (Delacourt ; Dournon ; Lis & Barbier ; Maloux ; Montreynaud-Pierron-Suzzoni, Pierron ; Rey-Chantreau ; Sommant) ; *Tout vient à point qui sait attendre* (Montreynaud-Pierron-Suzzoni, Rey-Chantreau) ; *Tout vient à point à qui peut attendre* (Dournon) ; *Tout vient à point qui peut attendre* (Montreynaud-Pierron-Suzzoni), *Tout vient à temps à qui sait attendre* (Dournon). Voici enfin un échantillon de ce qu'on peut trouver dans la presse (*Le Monde*) : *Tout vient à point qui sait attendre*, *Tout vient à point pour qui sait attendre*, *Tout vient à temps pour qui sait attendre*, *Tout arrive à qui sait attendre*. Et enfin les réseaux (SketchEngine/frTenTen12) : *Tout vient à point à qui sait attendre*, *Tout vient à point pour qui sait attendre*, *Tout vient à point à celui qui sait attendre*, *Tout vient à point à ceux qui savent attendre*, *Tout vient à point à qui veut attendre*, *Tout vient à temps à qui sait attendre*, *Tout vient à temps qui peut attendre*. Et il y a des cas beaucoup plus spectaculaires. On est donc très loin du proverbe vu comme affichant 'un figement complet' et ne tolérant 'aucun changement formel' comme l'affirme pourtant Schapira (1999 : 10).

c) Pire, les jugements d'acceptabilité se font soit par introspection, soit en interrogeant des autochtones. Or, nous avons pu vérifier par l'expérience que les autochtones qu'on interroge tendent à répondre non selon leur intuition, mais selon une grammaire normative qui a fixé depuis longtemps les règles de l'écrit. D'où très souvent un décalage entre une norme omniprésente mais implicite, et un usage effectif dont on ne soucie guère (cf. Blanche-Benveniste, 1989). Ainsi *vendre la mèche* est très généralement réputée non passivable, ainsi dans Lamiroy & Klein (2005 : 143), alors que cette locution apparaît au

passif dans certains corpus (cf. ci-dessus, 4.5.2.). De même pour *montrer patte blanche*, qui donne lieu à des exemples (non uniques) comme (*SketchEngine*, 13/03/2025):

« ... Comment veux-tu voir éclore un libre marché s'il est préalablement exigé, d'abord, que soit montrée patte blanche ... ».

Comment éviter d'interpréter ces données autrement que par de pseudo-statistiques ?

4.5.4. Bilan provisoire : figement, idiomatique et fixité

Au vu des remarques qui précédent, on comprend la difficulté que rencontre toute tentative de définition de la notion de figement. En fait, et comme il est fréquent en linguistique, les problèmes se cachent dans les nombreux *non-dits* de la discipline, en particulier lors de la mise en application. Dans le cas qui nous occupe, ces non-dits sont de trois types : a) le problème de l'origine et du statut des données présentées par le linguiste ; b) La contradiction entre la définition générale du figement et les critères évoqués à l'appui d'une hypothèse de figement ; c) La cohérence de la notion elle-même de figement.

Le premier point, que nous avons évoqué précédemment, pose la question de la validité des données utilisées, qui présente deux facettes intimement liées. La première concerne ce qu'il est convenu d'appeler les *corpus*. Toute utilisation de corpus dans une étude suppose en effet, nous l'avons dit, que ce corpus est exhaustif, i.e. couvre l'ensemble des manifestations du phénomène étudié. Or dans beaucoup d'études sur le figement, les données sont issues de la connaissance que le locuteur a ou croit avoir de sa langue, et qui ne sont bien souvent que le reflet d'un enseignement normatif de cette même langue, favorisant une certaine langue écrite et ignorant voire condamnant les productions orales (cf. Blanche-Benveniste 2010). Les exemples cités ne sont là le plus souvent que pour illustrer le bien-fondé d'une hypothèse. Un exemple-type de cette attitude est le statut des proverbes dans la plupart des recherches sur le figement. Ces malheureux proverbes sont réputés totalement figés, ainsi par Norrick (1985), Kleiber (1999), Schapira (*op. cit.*), Svensson (2004), Palma (2007), etc., ou à la rigueur plus figés que les locutions verbales. C'est ce que déclare Lamiroy (2021 : 249), en s'appuyant sur les trois exemples *Pierre qui roule n'amasse pas mousse*, *Un tiens vaut mieux que deux tu l'auras*, *Qui dort dîne*, sans justifier ce point de vue. Certes, la catégorie proverbiale comporte, comme toute catégorie linguistique, des éléments figés. Mais les recherches que nous avons menées sur le sujet montrent qu'en position de porter un jugement sur telle ou telle expression de la langue, les sujets natifs tendent très fortement à répondre selon la norme grammaticale, et non selon leur propre usage, nous l'avons dit *supra*. Ainsi, les exemples fournis par Lamiroy (*op. cit.*) sont ceux qui sont fournis par *Frantext*, mais si on consulte un corpus à vocation plus orale comme *SketchEngine*, des variations apparaissent : *A pierre qui roule, nulle mousse n'amasse*, *Pierre qui roule n'amasse jamais mousse*, *Pierre qui roule n'amasse pas de mousse*, *Pierre qui roule amasse peu de mousse*, *Pierre qui roule n'amasse toujours pas mousse*, *Pierre qui roule ne ramasse pas mousse* ; *Un bon tiens vaut mieux que deux tu l'auras* ; *Celui qui dort dîne*. Et ce cas n'est pas le plus extraordinaire que nous ayons rencontré. *Frantext* ne propose que les deux variantes *C'est dans les vieilles marmites qu'on fait les meilleures soupes* et *C'est dans les vieux*

chaudrons qu'on fait les meilleures soupes pour la version ‘officielle’ de *C'est dans les vieux pots qu'on fait les meilleures soupes*. Alors que SketchEngine (*esTenTen18*) en offre plusieurs dizaines ! (Anscombe 2022a). On pourrait penser que chaque version est restreinte à une communauté différente, et qu'il ne s'agit donc pas de variantes, mais de versions distinctes que rien n'empêche alors de considérer comme figées 'séparément'. Or pour une version donnée d'une parémie, il n'est pas rare de trouver chez un même usager plusieurs autres versions de la même, ce qui réduit cet argument à néant. En voici un exemple, celui de *Qui vit par l'épée, périra par l'épée*, version standard d'origine biblique. Il donne lieu à quatre variantes chez l'auteur de romans d'espionnage Gérard de Villiers, à savoir : *Qui frappe par l'épée, périra par l'épée*, *Celui qui a frappé par l'épée, périra par l'épée* ; *Celui qui frappe par l'épée, périra par l'épée* ; *Celui qui se sert de l'épée, périra par l'épée*³². Et il ne s'agit pas là d'un cas unique : les variantes sont fréquentes en particulier chez les auteurs où les dialogues abondent. On pourrait penser que l'exemple que nous venons d'exposer provient en fait d'une littérature d'un niveau de langue médiocre, proche donc d'un oral qui a toujours été considéré comme une langue défectueuse. Il n'en est rien : l'exemple suivant concerne l'auteur espagnol contemporain Arturo Pérez-Reverte, célèbre pour ses romans d'aventure. Il est peu susceptible d'une maîtrise discutable de la langue, étant donné qu'il est membre de la *Real Academia Española*, l'équivalent espagnol de l'Académie française. Il nous propose le très courant proverbe *A quien madruga, Dios le ayuda* (litt. 'A qui se lève tôt, Dieu l'aide'), correspondant au français *L'avenir appartient à ceux qui se lèvent tôt*, mais aussi trois variantes distinctes, à savoir : *Al que madruga Dios le ayuda* ; *A quien madruga, etcétera* ; *Al que madruga, Dios lo ayuda*³³. Ce phénomène de variantes parémiologiques semble donc être dû au statut linguistique général des proverbes, et non pas lié à une langue particulière. Il convient par conséquent d'être prudent à l'heure d'émettre des jugements d'acceptabilité, et de ne pas se limiter à des intuitions concernant quelques rares exemples hors contexte, quand ils ne sont pas extraits de dictionnaires au contenu discutable. Nous avons examiné en 4.5.2 le critère de passivation sur le cas de *pousser les hauts cris*, qui aboutit aux mêmes conclusions. Voici maintenant le cas d'un autre critère, celui des reprises anaphoriques des termes d'une expression figée. Toujours selon Lamiroy (*op. cit.* : 256), « ... les compléments figés sont difficilement pronominalisables, ayant perdu, entièrement ou partiellement, leur sens référentiel ... ». Et de proposer la paire :

(19) Max a repris du poil de la bête/ ? Max **en** a repris, du poil de la bête.

qui, là encore, repose sur la seule intuition de l'auteur à propos de deux exemples sortis de tout contexte. Or l'exemple suivant :

³² Respectivement : *La bataille des S6300-T1*, Ed. Plon, Paris, 2009 : 88 ; *Les fantômes de Lockerbie*, Ed. Plon, Paris, 2013 : 33 ; *Le trésor de Saddam* : 2, Ed. Plon, Paris, 2013 : 258 ; *Vengeance romaine*, Ed. Plon, Paris, 1998 : 182.

³³ Respectivement: *La tabla de Flandes*, Alfaguara, Madrid, 2005: 534.; *El caballero del jubón amarillo*, Punto de lectura, Madrid, 2005: 74; *La piel del tambor*, Ed. De Bolsillo, Barcelona, 2003: 475; *El sol de Breda*, Punto de lectura, Madrid, 2003: 231.

- (20) Depuis qu'il est sorti de l'hôpital, je peux te dire qu'il en a repris, Max, du poil de la bête !

a été jugé parfait par tous les francophones interrogés. On notera que ce dernier exemple restitue un contexte et un cotexte, ainsi qu'un minimum de ligne prosodique. Concluons : beaucoup de travaux sur le figement se fondent non sur un ou des corpus, mais sur des unités lexicales en petit nombre et hors-contexte, assortis de jugement d'acceptabilité mal contrôlés. Nous pensons, à l'inverse, que toute étude de figement doit se faire à partir d'un corpus au sens scientifique du terme, *i.e.* un ensemble construit de données de façon à satisfaire les préalables déjà évoqués d'exhaustivité, de pertinence et d'homogénéité. Ce qui met au centre du problème la contribution des corpus à l'analyse linguistique, et questionne au passage l'usage débridé qui est parfois fait des 'grands corpus', souvent présentés comme le *nec plus ultra* de la scientificité et donc la panacée aux insuffisances théoriques. On pourra se reporter à Balibar-Mrabti (2020) sur le sujet.

Venons-en maintenant au problème de la définition générale du figement et des critères qui sont habituellement proposés comme reflétant ce figement.

Rappelons brièvement ce qui a été dit précédemment : toutes les études sur le figement s'appuient sur une définition générale du figement et sur une série de critères. La définition générale est la non-compositionnalité : une expression figée est non-compositionnelle, *i.e.* que la connaissance des valeurs sémantiques des lexèmes qui la composent ne permet cependant pas d'en déduire le sens global. Il est en germe dans le critère B₅ de Bally ('l'oubli du sens des éléments'), et figure surtout chez Jespersen, pour qui le sens du tout n'est pas déductible du sens des parties. Cette non-compositionnalité est supposée se refléter dans trois catégories de phénomènes qui lui servent de critères :

- a) le figement référentiel, ou non actualisation des éléments (proche du critère B₃ de Bally),
- b) le figement transformationnel, ou blocage de certaines manipulations de la structure de surface (proche du critère B₂) ; et enfin,
- c) Le figement sémantique, ou encore opacité sémantique (variante du critère B₅).

Il y a certes des expressions réputées figées et qui satisfont largement à l'ensemble de ces critères. Considérons par exemple la locution verbale *mettre la main à la pâte*. La consultation de plusieurs corpus ne nous a pas permis de découvrir quelque variation que ce soit pour cette locution, hormis un cas de *remettre la main à la pâte*, qui peut d'ailleurs être interprété comme *re-[mettre la main à la pâte]* et ne contrevenant donc pas aux critères de figement transformationnel. Le passif semble en particulier impossible. Par ailleurs, la non-compositionnalité est acquise. Ce cas idyllique semble malheureusement exceptionnel. Dans la même série avec *mettre*, considérons en effet *mettre du cœur à l'ouvrage* : la recherche dans divers corpus a confirmé l'impression que le passif *Du cœur avait été mis à l'ouvrage* est difficile. Par ailleurs, il y a une certaine opacité, qui n'est pas totale, du moins dans le domaine indo-européen. Là s'arrête le côté figé de la locution. D'une part, le verbe *mettre* peut être remplacé par *remettre, donner, avoir* ; d'autre part, on trouve dans les corpus *mettre tout son cœur à l'ouvrage* (*Frantext*) et ...*en y mettant chacun le même cœur à l'ouvrage* (*SketchEngine*). Il y a plus : ainsi, la locution *mettre le*

feu aux poudres admet aussi la variation *remettre, fouter, flanquer* pour *mettre*. Mais *SeketchEngine* fournit 70 cas de passivation du type de *le feu a été mis aux poudres*. Pourtant, les deux locutions *mettre du cœur à l'ouvrage* et *mettre le feu aux poudres* ont la même structure de surface *mettre + SN + Prép + SN*, et sont toutes deux non-compositionnelles. Toutes deux admettent certaines variations au niveau du verbe, mais bizarrement, une seule des deux se passive. On découvre sur cet exemple simple un des problèmes de la *doxa* concernant le figement. Il ne semble pas y avoir de relation systématique entre la non-compositionnalité d'une part, et la réponse aux différents critères évoqués à ce propos, y compris pour des structures proches, d'où certaines hésitations sur des cas concrets (Vaguer, 2005 : 231 *sq.*). Entre autres, pourquoi *mettre du cœur à l'ouvrage* admet une certaine actualisation du complément d'objet, et non *mettre la main à la pâte*? Pourquoi aucune des deux ne se passive, au contraire de la locution *mettre le feu aux poudres*? Les locutions verbales ne sont pas les seuls éléments à poser ces problèmes, les proverbes (exemple d'unités polylexicales phrastiques) présentent les mêmes. Ainsi, on ne peut rien changer à la parémie *Abondance de biens ne nuit pas*³⁴, qui est très généralement considérée comme totalement figée. Pourtant, elle est compositionnelle, de même que *Les absents ont toujours tort ; Loin des yeux, loin du cœur ; Le cœur a ses raisons que la raison ignore ; Chaque chose en son temps*, etc., également considérées comme des proverbes et donc en principe figées. Il faudrait donc admettre qu'il y a des idiomes qui sont compositionnels tout en étant figés, ce qui peut laisser perplexe ... Signalons enfin le domaine des locutions prépositives, déjà signalé par G. Gross (2012 : 140 *sq.*) comme marqué par l'irrégularité, et qui pose les mêmes questions : certaines admettent des variations d'intensité par introduction d'un adjectif : *au milieu de/au beau milieu de, au dam de/au grand dam de, au fond de/au fin fond de*, etc. Pour d'autres, il faut pour ce faire changer également la préposition : *au milieu de/en plein milieu de, au cœur de/en plein cœur de, à la figure de/en pleine figure de, au travers de/en plein travers de*, etc. Tous ces phénomènes, dont l'explication ne semble pas aller de soi, pointent tous vers une même direction : la question de la cohérence de la notion de figement.

En effet, la définition générale du figement par la non-compositionnalité, représentée et prolongée par les critères que nous avons évoqués à plusieurs reprises, recouvre une hypothèse extrêmement forte, déjà pointée du doigt, et qui est que l'ensemble des expressions figées forme un tout homogène, caractérisable par l'ensemble bien défini de ces critères. Ruwet (*op.cit.*) a été un des tout premiers à remarquer que nous sommes incapables de prédire, pour un idiom donné, quelles 'transformations' syntaxiques spécifiques vont le caractériser comme tel. En atteste le fait que deux expressions possédant des structures superficielles semblables ne permet pas d'en déduire des comportements identiques par rapport aux critères usuels de figement. Nous avons évoqué *supra* les exemples de *mettre le feu aux poudres* versus *mettre du cœur à l'ouvrage* ; de *mettre du cœur à l'ouvrage* versus *mettre la main à la pâte*, parmi d'autres. Ces expressions ont des structures quasi-identiques ; plus, leur sens sont proches. Et nous avons vu (cf. 3.2. *supra*) que pour Nunberg, Sag & Wasow (*op. cit.*), il conviendrait de

³⁴ Certains locuteurs admettent la variation *Abondance de biens n'a jamais nui*.

distinguer, pour un idiome donné, la *conventionnalité* – i.e. la non-prédicibilité du sens à partir du sens que possèdent les composants dans une combinaison libre, répartie sur l'ensemble des composants, et le *sens figuré*, qui repose sur une métaphore globale. Par exemple, *casser sa pipe* serait métaphorique, mais non *vendre la mèche*. Pour *casser la pipe*, le cas est relativement clair. L'origine est argotique, et les sens des deux composants sont bien attestés : *casse-pipe*, *casse-cou*, *casser la figure*, *les gueules cassées*, pour *casser*. *Pipe* a eu en argot, jusqu'au début du XX^e siècle, les sens de 'figure' et aussi de 'tuyau, gorge' (Sainéan, 1920). Il s'agit donc bien d'une métaphore, *grosso modo* 'couper le souffle'. Pour *vendre la mèche*, il s'agit en fait d'une déformation de *éventer la mèche*, qui signifiait, selon Oudin (1640 : s.v. *mesche*) 'descouvrir la mesche', i.e. descouvrir la malice ou finesse, vulg. '. Il n'y a pas métaphore, à l'instar du correspondant anglais *to spill the beans*, les errements de l'évolution diachronique ayant masqué un éventuel sens littéral. Certes, si la critique de Nunberg, Sag & Wasow n'est pas dénuée de fondement, la distinction est souvent difficile à faire sur des cas concrets. Elle met cependant le doigt sur un aspect assez largement méconnu dans les études sur le figement, nous l'avons déjà dit, et qui est la dimension diachronique. Comme le montrent en effet certains des exemples que nous venons d'évoquer, le statut d'expression figée résulte d'un processus plus ou moins long dans le temps. Ce processus ne s'installe pas au même moment pour toutes les expressions, d'où il résulte que dans une coupe synchronique donnée, deux expressions particulières n'en sont pas nécessairement au même stade de leur processus de figement, et sont donc susceptibles de ne pas satisfaire aux mêmes critères. Cette même dimension diachronique est à l'origine du fait qu'une métaphore à l'origine de la formation de l'expression a pu être occultée au cours de l'évolution diachronique, et donc disparaître. Ainsi l'exemple cité *supra* de *éventer la mèche*, non métaphorique, face à l'expression cette fois métaphorique *bête comme ses pieds*, les pieds étant traditionnellement reliés à la médiocrité – *penser avec les pieds*, *travailler avec les pieds*, métaphore qui semble d'origine romane. L'espagnol familier *con los pies* signifie 'maladroitement, mal', ainsi dans *escribir con los pies* 'écrire mal'. Il est probable qu'une partie des disparités constatées parmi les expressions réputées figées ait pour origine des parcours diachroniques différents. Dernier point enfin, et non des moindres, l'idée même du figement. Elle repose en effet, rappelons-le une nouvelle fois, sur le critère de non-compositionnalité, les divers critères spécifiques évoqués – syntaxiques, sémantiques, référentiels – étant supposés n'être que des illustrations particulières des effets de ce figement. De ce point de vue, la catégorie proverbiale, dans la mesure où elle est très largement considérée comme composée d'éléments figés, remet en question la nature même du figement. En effet, une sous-catégorie importante de proverbes exhibe une fixité formelle indiscutable, ainsi : *Abondance de biens ne nuit pas*, *Mieux vaut tard que jamais*, *L'argent ne fait pas le bonheur*, *A l'impossible, nul n'est tenu*, *C'est l'intention qui compte*, *Chaque âge a ses plaisirs*, *La colère est mauvaise conseillère*, *Le temps perdu ne se rattrape jamais*, *On n'est jamais trahi que par les siens*, *On risque de tout perdre en voulant trop gagner*, etc. Ces mêmes proverbes ont également la particularité d'être compositionnels, d'où le problème. Si l'on veut conserver le critère de non-compositionnalité comme base du figement, on doit alors admettre que : a) tous les proverbes ne sont pas figés en ce sens ; b) La fixité (ou une certaine fixité) de la forme n'a pas nécessairement à voir avec le figement ainsi défini.

Dans la foulée, on remarque que certaines locutions prépositives comme *au milieu de*, *au mépris de*, *au fond de*, *par les soins de*, *face à face avec*, etc., présentent le même problème que les formes parémiques. Tout en présentant une certaine rigidité de forme – qui les fait souvent considérer comme figées³⁵ – elles sont visiblement compositionnelles. Si donc le figement entraîne souvent une fixité formelle plus ou moins grande, l'inverse n'est pas vrai. Concluons : les hiatus que nous avons rencontrés entre définition non-compositionnelle et définition par critères sémantiques + référentiels + transformationnels proviennent en fait d'une confusion entre deux types de phénomènes : a) les phénomènes de figement ; b) l'existence de structures présentant une certaine fixité. C'est à la thèse de cette dichotomie que nous allons désormais nous consacrer.

5. FIGEMENT VERSUS CONSTRUCTIONS, PATTERNS, SCHÉMAS LEXICAUX ET MATRICES LEXICALES

Ce chapitre est consacré à des approches récentes en linguistique, celles qui se placent du point de vue de *patrons* – quelles que soient les hypothèses retenues à leur sujet – qui sous-tendent la formation non seulement des expressions 'ordinaires', mais aussi celle des formations dites phraséologiques. Il est bien sûr impossible d'examiner toutes les variations autour d'un tel thème, et nous nous contenterons donc, dans le fil de cet exposé, de situer les principales approches d'un point de vue théorique, en les opposant en particulier à la position de la grammaire chomskyenne, référence habituelle des travaux sur la phraséologie. Pour des études plus détaillées, on pourra se reporter à Goldberg (1995), Mathieu (2003), Legallois & François (2006), François (2008), entre (beaucoup d') autres. Pour illustrer le type de problème que constitue l'étude d'un phénomène linguistique selon une telle optique, nous évoquerons pour terminer l'analyse complète de trois 'patrons' qui relèvent selon nous d'une telle approche, à savoir : les proverbes (Anscombe, 2011), la classe des expressions du type *faire ses adieux* (Anscombe 2022b, 2022c), et enfin les 'noms de bruit' suffixés en *-ido* en espagnol (Anscombe 2024). Ces trois études font partie des rares travaux sur le figement abordant non des expressions isolées, mais des catégories d'expressions.

5.1. Les grammaires de construction

La plupart des théories sémantiques calculent la valeur sémantique d'une phrase comme issue des valeurs sémantiques des mots qui la composent, selon un processus plus ou moins compliqué. C'est ce que l'on appelle les *théories compositionnelles du sens*. Dans le cadre générativiste par exemple de Chomsky (1989), et résumé très grossièrement, la description d'un mot comporte des renseignements sur les contraintes syntaxiques et sémantiques qu'il impose pour pouvoir être la tête (*the head*) d'un syntagme. C'est ce que l'on appelle sa *structure argumentale*, et ce principe est appelé *principe de projection lexicale*. Ces syntagmes sont à leur tour assemblés pour former des phrases. Notons que dans un tel cadre, les règles d'assemblage des mots pour former des syntagmes sont purement syntaxiques : les entités sémantiques comme les événements ou les propriétés ne sont exprimées qu'au niveau des mots, ce qui signifie entre autres choses que, ces

³⁵ Cf. cependant certains doutes sur le sujet dans Gross (1996 : 126-127).

règles combinatoires étant syntaxiques, les phrases ont bel et bien un sens, mais que les structures de phrases n'en ont pas. Cette représentation est en fait en accord avec le programme minimaliste³⁶, pour lequel les cadres syntaxiques représentant une langue n'ont pas de signification propre. Les structures d'une langue sont le résultat de l'interaction de principes généraux (entendez *universels*) et de paramètres propres à chaque langue (Chomsky, 1995 =1997 : 117) :

A language is not, then, a system of rules, but a set of specifications for parameters in an invariant system of principles of Universal Grammar (UG); and traditional grammatical constructions are perhaps best regarded as taxonomic epiphenomena, collections of structures with properties resulting from the interaction of fixed principles with parameters set one or another way. There remains a derivative sense in which a language L is a “rule system” of a kind: namely, the rules of L are the principles of UG as parameterized for L.

C'est à cette coupure entre syntaxe et sémantique que s'opposent les *grammaires de construction(s)* (en abrégé CxG), puisqu'elles partent du principe que les structures syntaxiques elles-mêmes sont dotées de signification. De façon plus précise, les CxG font l'hypothèse d'une structure fondamentalement phraséologique de la langue. Les contraintes qui régissent forme et interprétation ne peuvent être expliquées que comme le produit de *constructions grammaticales*, en entendant par là un réseau d'appariements (forme, sens) où les deux pôles sont reliés de façon souvent complexe, et à la productivité (*entrenchment*) variable³⁷. La grammaire est dans cette perspective, vue comme l'inventaire raisonné de ces appariements. Il s'agit donc bel et bien d'une contestation de toute théorie générative qui, à l'instar de celle de Chomsky, verrait une opposition irréductible entre le lexique (formé d'unités ayant un sens) et la grammaire (dont les structures n'ont pas de sens). Cette contestation s'appuie sur quatre points fondamentaux :

1. Les structures linguistiques sont au moins partiellement idiomatiques. C'est l'axiome de départ des CxG, qui est l'hypothèse ci-dessus d'une structure fondamentalement phraséologique de la langue. Pour les CxG, les constructions dites figées (ou phraséologiques, ou idiomatiques, etc.) font partie d'une langue au même titre que les constructions libres, et de ce point de vue, les CxG ne distinguent pas des constructions normales et des constructions anormales, comme c'était très généralement le cas, et pas seulement dans le cadre génératifiste.
2. Une conséquence immédiate de cette hypothèse est qu'un schéma combinatoire peut posséder une valeur sémantique intrinsèque, non-réductible donc à une logique

³⁶ Pour le programme minimaliste, cf. Chomsky (1989, 1997) ; également Sag & Wasow (1999), et la très claire présentation de Hornstein, Nunes & Grohmann (2005). La théorie des *Principes et paramètres* est exposée dans Chomsky (1997) par H. Lasnik (11-116).

³⁷ Ce point avait été explicitement noté dès Ruwet (1983 33) : « ... Plusieurs linguistes [...] l'avaient pressenti, mais c'est Nunberg (1978 : 117 *sq.*) qui l'a le mieux montré : le comportement des Exis et spécialement leur caractère plus ou moins figé, ne peut se comprendre ni par leur forme seule, ni par leur sens seul, mais par la relation entre leur forme et leur sens - ce qui revient à dire que les Exis sont, à des degrés divers, analysables, et que leur sens littéral (ou le sens littéral de leurs constituants) reste presque toujours pertinent ... ».

combinatoire fondée sur le seul principe de projection. Par conséquent, de ce point de vue, les structures sous-jacentes sont sémantiques.

3. Il n'y a pas de synonymie entre structure sous-jacente et structure dérivée. En d'autres termes, les *patterns* ne peuvent être synonymes : une différence dans les formes syntaxiques entraîne de façon quasi-automatique une différence de sens.

4. Enfin, la structure des formes telles qu'elles apparaissent en surface – et en particulier des patterns, est motivée. Ce qui autorise une construction, un pattern, à renvoyer à une valeur sémantique globale qui soit distincte de la simple somme des sens lexicaux de ses composants.

Dans une telle optique, la grammaire d'une langue est une grammaire de *constructions*³⁸, i.e. de patrons théoriques (au sens des appariements évoqués ci-dessus) qui servent à former les éléments de la langue, du mot lexical jusqu'aux phrases en passant par les syntagmes. Voici la définition du terme de base *construction* telle qu'elle apparaît dans Goldberg & Jackendoff (2004 : 532-533) :

THE CONSTRUCTIONAL VIEW

- a. There is a cline of grammatical phenomena from the totally general to the totally idiosyncratic.
- b. Everything on this cline is to be stated in a common format, from the most particular, such as individual words, to the most general, such as principles for verb position, with many subregularities in between. That is, there is no principled divide between 'lexicon' and 'rules'.
- c. At the level of phrasal syntax, pieces of syntax connected to meaning in a conventionalized and partially idiosyncratic way are captured by CONSTRUCTIONS.

Ces grammaires de construction(s) ont été popularisées par Goldberg (1995), mais leur apparition reste redéivable à certains travaux antérieurs. Parmi les plus cités figure en particulier Fillmore *et alii* (1988). On trouvera également des développements ultérieurs dans Kay & Fillmore (1997), Goldberg & Jackendoff (2004), Goldberg (2006), parmi beaucoup d'autres. Par ailleurs, la notion de construction a été reprise de différentes façons par divers auteurs, à part Goldberg qui représente le courant 'cognitif', ainsi Croft (2001) pour la version dite 'radicale'. On pourra consulter François (2008) pour les divers développements autour de cette notion, ainsi que Legallois & Patard (2017). Enfin, Legallois (2018) présente un exposé pédagogique complet sur le phénomène des CxG.

Voici, grossièrement résumée, la façon dont Goldberg (1996 : 199-218) utilise la notion de construction pour rendre compte du comportement de *way* 'route, chemin, manière' dans des exemples comme :

³⁸ Pour éviter une possible confusion avec d'autres théories, comme les *grammaires de patterns*, nous adopterons le terme de *construction* pour la théorie de Goldberg et apparentées. Le terme anglais correspondant est *construct*, mais nous n'avons pas voulu ajouter un nouvel anglicisme à un vocabulaire linguistique déjà chargé d'emprunts plus ou moins heureux.

- (21) *Frank dug his way out of the prison.*
 ('Frank a creusé sa *way* pour sortir de prison' = Frank s'est évadé de prison en creusant un tunnel).
- (22) *Frank found his way to New York.*
 ('Frank a trouvé sa *way* vers New-York' = Frank a trouvé (et pris) le chemin de New-York).

Comme le remarque l'auteur, de telles phrases présentent le référent du sujet comme se déplaçant le long du trajet désigné par le syntagme prépositionnel. La comparaison avec:

- (23) *Frank dug his escape route out of prison.*
 ('Frank a creusé un chemin pour s'évader de prison' = Frank a creusé un tunnel pour s'évader de prison).
- (24) *Frank found a way to New York.*
 ('Frank a trouvé une *way* pour New-York' = Frank a trouvé le moyen d'aller à New-York)

montre que

a) dans (23) et (24), le syntagme prépositionnel modifie l'objet direct. En effet, ni (23) ni (24) n'impliquent de mouvement, comme le montrent les enchaînements avec *but he hasn't gone yet* (= 'il n'est pas encore parti') :

- (23a) Frank dug his escape route out of prison, but he hasn't gone yet.
 (24a) Frank found a way to New York, but he hasn't gone yet.

b) À l'inverse, (21) et (22) impliquent un tel mouvement, bien qu'il ne soit impliqué par aucun des composants lexicaux :

- (21a) *Frank dug his way out of prison, but he hasn't gone yet.
 (22a) *Frank found his way to New York, but he hasn't gone yet.

c) Étant donné que l'interprétation de ce type de construction en *-way* n'est pas totalement prédictible à partir du sens de ses composants lexicaux, Goldberg en propose une analyse en termes de construction, nommément le schéma de la 'Construction way':

SUJET+ VERBE [Possessif + *way*] + OBJET Locatif

On trouvera dans Goldberg (*op. cit.*) une discussion détaillée de ces divers points et des comparaisons avec des combinaisons en apparence semblables. Goldberg (*op. cit.*: 200) reprend l'interprétation de Jackendoff (1990: 221) selon laquelle:

«... in a sense, the *way*-construction can be thought of as a kind of 'constructional idiom,' a specialized syntactic form with an idiomatic meaning, marked by the noun *way* ...». On notera les deux points importants dans cette déclaration : les constructions sont un doublet insécable (forme, sens), et cet appariement – qui seul permet d'accéder au sens du tout – est marqué par la présence du nom *way*³⁹.

³⁹ On trouvera une critique partielle de cette analyse des 'way-constructions' dans Luzondo Oyón (2013).

Si on tente de résumer la philosophie linguistique sous-jacente à la notion de construction selon Goldberg, plusieurs caractéristiques s'en détachent :

a) Une construction est fondamentalement, comme nous l'avons dit *supra*, une paire de type (F, S), où F est une forme et S un sens. Cette paire est conventionnalisée, à savoir que l'appariement entre F et S est insécable. Comme l'exprime déjà Lakoff (1987 : 467) :

Each construction will be a form-meaning pair (F, M), where F is a set of conditions on syntactic and phonological form and M is a set of conditions on meaning and use.

b) Ce type de correspondance entre forme et sens relève des modèles symboliques ayant une réalité cognitive (*op. cit.* : 467) :

An even more general theory of symbolic models would account for the understanding of all sorts of form-meaning correspondences that have a cognitive reality.

Il s'agit donc d'une théorie de type *usage-based* (Langacker 1987 : 46):

Cognitive grammarians regard statements as the matrix from which general statements (rules) are extracted.

c) La forme d'une construction n'est pas dérivable d'autres formes, même si elle est apparentée à une structure plus générale (elle est 'motivée'). En particulier, elle n'est pas complètement compositionnelle⁴⁰.

d) Dans la mesure où la notion même de construction rejette la séparation entre syntaxe et sémantique d'une part, et entre lexique et syntaxe d'autre part, les constructions peuvent concerner tous les domaines de la langue : des unités lexicales (ainsi le mot *way* abordé *supra*), la morphologie dérivationnelle ou flexionnelle, comme le *s* du pluriel anglais dans certains items lexicaux (Goldberg, 2006 : 218 *sq.*), des suites polylexicales conventionnalisés comme *Much ado for nothing*, des phrasèmes figés comme *to kick the bucket*, etc.

e) Enfin, *last but not least*, les constructions ne relèvent pas de phénomènes de figement qui au terme d'un processus plus ou moins long, engendreraient une combinaison donnée hors des règles de la langue considérée, tant syntaxiques et morphologiques que sémantiques. Il s'agit en fait de phénomènes de *fixité* qui sont précisément liés au fait que dans une construction, forme et sens se contraignent réciproquement, et que c'est ce lien qui marque la construction comme telle. Nous aurons l'occasion de revenir sur ce point lors de la conclusion de ce chapitre.

Certaines thèses de Goldberg ont fait l'objet de critiques, en particulier celle exposée en b) *supra*. A chercher à tout prix une 'réalité cognitive' derrière les constructions, il arrive en effet qu'on 'force' en quelque sorte cette même réalité. C'est le cas de l'explication du -*s* pluriel final dans des mots – les 'lower trunk-wear' – comme *pants, shorts, knickers, kulos, leggings, stockings, trousers, khakis*, etc (Goldberg 2006: 218-19):

⁴⁰ Comme le note François (2008), cette contrainte figure chez Goldberg, mais est rejetée par Langacker (1987). On trouvera chez François (2008) une comparaison détaillée des différentes théories constructionnelles.

All English words that refer to lower trunk-wear are grammatically plural, e.g. *pants*, *shorts*, *knickers*, *kulots*, *leggings*, *stockings*, *trousers*, *khakis* (Williams, 1994). As Langacker (1987a: 47) points out, this type of grammatical plurality is motivated by the fact that the referents involved all have bipartite structure. Lower-trunk-wear all have two parts, one for each leg. Notice skirt and wrap are non-bipartite and as expected are also not grammatically plural.

The general, productive plural construction and the lower-trunk-wear plural construction share the same form and have related meaning. Because the formal similarity indicates a semantic relationship, the lower-trunk-wear construction is motivated.

On note immédiatement que les pluriels mentionnés ne possèdent très généralement pas de version au singulier. Et si tel est le cas, la version au singulier ne possède pas le même sens, comme on peut le voir sur l'exemple français *lunette/lunettes*, ou encore l'anglais *brain* ('cerveau') / *brains* ('intellect'). Ce qui les distingue de cas 'réguliers' comme *chaussette/chaussettes* ou encore *shirt/shirts*. Un premier problème est le choix des termes relatifs aux 'lower-trunk-wear', sous prétexte que ces vêtements sont caractérisés par le fait qu'ils ont une structure bipartite : ils sont composés de deux parties, une pour chaque jambe. Or ce choix et le raisonnement qui l'explique paraissent quelque peu *ad hoc* : une chemise possède aussi une structure bipartite en ce sens, à savoir une manche pour chaque bras, et se nomme en anglais *shirt* et non *shirts*. Par ailleurs, le vocabulaire des lower-trunk-wear ne sont pas les seuls à posséder cette propriété : figurent parmi les lexèmes possédant cette propriété des noms d'outils ou d'objets : *scissors*, *spectacles*, *nippers*, *tongs*, *glasses*, *pliers*, etc. Or bizarrement, *hair grip*, *clothes pin* et (*paper*) *clip* – respectivement 'pince à cheveux', 'pince à linge' et 'trombone' (lit. 'pince à papier'), ne se présentent pas sous cette forme plurielle. Il y a plus : le corps humain présente, par rapport aux jambes, une structure bipartite au sens de Goldberg ; pourtant, *leg* existe bel et bien au singulier. Enfin, *bermuda* et *bathing suit* ne portent pas le *s* de pluriel, ce qui peut surprendre. On voit où gît le problème : le besoin d'une 'réalité cognitive' a fait choisir un domaine où le concept de construction semblait pouvoir s'appliquer sans problème. Nous pensons que le point faible de l'hypothèse de Goldberg réside en ce que le *s* étudié a été qualifié de *s pluriel* sans autre forme de procès. Or ce *s* sert non seulement en anglais, mais dans l'ensemble des langues indo-européennes, à de multiples fonctions⁴¹. En particulier, à désigner des ensembles pluriels d'éléments liés entre eux, mais non dénommables par le vocable au singulier : d'où, en français, l'existence de pluriels comme *fiançailles*, *funérailles*, *agapes*, *festivités*, *retrouvailles*, *alentours*, *pourparlers*, etc. Il convient donc de les distinguer des pluriels en *-s* possédant un singulier destiné à dénommer des éléments semblables. Il nous semble par conséquent qu'une étude détaillée du fonctionnement du pluriel devrait être menée afin d'éviter toute hypothèse risquant de se révéler *ad hoc*. D'autant plus que la présence de cas similaires dans d'autres langues et sur des mots analogues (ainsi angl. *glasses*, fr. *lunettes*, esp. *gafas*, ital. *occhiali*, etc. ; angl. *pants*, fr. *pantalons*, esp. *pantalones*, ital. *pantalon*) tendrait à prouver que les usages, étant similaires dans ces langues, pointent vers une même réalité cognitive, indépendamment des environnements des langues considérées.

⁴¹ En particulier le *-s* dit adverbial, dont l'histoire mouvementée devrait inciter à la prudence. Voir à ce sujet Corbett (2019) et Anscombe (2022b).

Dernier point enfin : il nous semble que, tout en prônant l'absence de frontière entre expressions idiomatiques et expressions libres, Goldberg privilégie comme constructions des cas apparemment non (totalement) libres, comme *ce s pluriel*, qui de plus, est une combinatoire peu productive⁴². Pourquoi ne pas avoir considéré le pluriel usuel comme une construction à part entière ? On touche là un point qui n'est que peu évoqué chez Goldberg, qui est que ce qui caractérise une construction semble être la non-prédicibilité, elle-même liée au caractère conventionnel qui la caractérise, ici un *s pluriel* qui n'en est pas vraiment un. Et c'est cette conventionnalité qui fait choisir par exemple le *s pluriel* de *trousers*⁴³, plutôt que d'envisager le pluriel ordinaire.

5.2. La grammaire de pattern⁴⁴

Il s'agit d'une approche qui met au centre de la structure linguistique tant syntaxique que sémantique la notion de *collocation*. L'origine de cet intérêt pour les collocations est à chercher dans un aspect spécifique de la didactique britannique dans le domaine des langues étrangères : la maîtrise d'une langue étrangère passerait en particulier par celle des collocations. L'idée de base, que l'on trouve déjà chez Sinclair (1991), puis Sinclair (2004), est que les phénomènes linguistiques ne peuvent être appréhendés que dans leur contexte d'emploi. Ce n'est qu'à ce niveau – et non au travers d'intuitions plus ou moins contrôlées à propos d'exemples inventés – que se manifestent les faits grammaticaux. Dans une telle optique, il n'y a donc pas de clivage entre linguistique grammaticale et linguistique textuelle. Enfin, cette analyse linguistique s'opère à travers l'examen de corpus. On notera à ce propos qu'une telle approche n'a été rendue possible que par les progrès du traitement informatique de corpus, et était inenvisageable à l'époque de l'apparition des thèses générativistes. Le point de départ des travaux de Sinclair a été la constatation que, loin d'être des exceptions marginales, les 'patterns' collocationnels apparaissaient fréquemment dans les concordances lexicales dès lors qu'on passait au niveau du corpus. Il y a plus (Sinclair 2004: 3) :

Idiomaticity, far from being a marginal aspect of language, seemed to be ubiquitous and at least as significant as syntax in the construction of meaning (Sinclair 1991: 112). Fixed, repeated strings were ubiquitous in Sinclair's data, and, most markedly, in collocations that involve the most frequent everyday words. This led Sinclair to assert the existence of a tight bond between form and meaning, between sense and structure, and to the conclusion that collocations and idiomatic but very frequently occurring combinations were the real glue that held texts together.

⁴² Cette productivité faible a pour effet, dans certains cas, de provoquer un retour au singulier, par disparition du 's pluriel'. Ainsi le français *pantalon* ou *jean*, l'espagnol *pantalón* et *tejano*, l'italien *pantalone*, où la forme au singulier tend à supplanter la forme avec *s pluriel*. Signalons également le français *pince* et l'espagnol *pinza* en lieu et place respectivement de *pinces* et *pinzas*.

⁴³ Nous rejoignons ici Nunberg, Wasow & Sag (1994), pour qui c'est la conventionnalité qui fonde l'idiomaticité (cf. *supra*, 3.2), et non la non-compositionnalité. Cf. une position identique chez Ruwet (1983 : 33-34 ; et ici, note 82).

⁴⁴ Nous avons conservé sciemment le terme anglais *pattern*, pour marquer la différence d'avec les grammaires de construction, les deux approches présentant des traitements différents des phénomènes phraséologiques

Pour Sinclair, les collocations sont donc à la fois un principe organisateur de la langue – les mots sont définis par leurs collocations, et un principe organisateur du texte/du discours, dans la mesure où ces collocations sont le ciment ('the glue') qui assure la cohésion de ce texte/discours. Cette approche, originale s'il en fut, inverse ainsi les rôles respectifs de la grammaire et du lexique. Comme le dit Hoey, un des co-fondateurs du mouvement (2005: 1):

In this book I want to argue for a new theory of the lexicon, which amounts to a new theory of language. The theory reverses the roles of lexis and grammar, arguing that lexis is complexly and systematically structured and that grammar is an outcome of this lexical structure. The theory grew out of an increasing awareness that traditional views of the vocabulary of English were out of kilter with the facts about lexical items that are routinely being thrown up by corpus investigations of text. What began as an attempt to account for collocation turned into an exploration of grammatical, semantic, sociolinguistic and text-linguistic phenomena.

La phraséologie devient ainsi *le principe organisateur du discours*, et non plus un domaine d'anomalies. Ce rôle central de la phraséologie apparaît dans un des principes formulés par Sinclair, à savoir l'*idiom principle*, selon lequel le choix de structures plus ou moins 'préfabriquées' constitue un choix simple pour un locuteur, choix qui prime sur celui d'une structure 'libre' (Sinclair, 1991 : 109-114). En effet, les locuteurs tendent à privilégier les emplois les plus fréquents, dont il se trouve que ce sont généralement les plus idiomatiques.

Peu connue en Europe, la *Pattern Grammar*, qui relève de la *linguistique contextualiste*⁴⁵, est une contribution principalement britannique, dont le point de départ est l'ouvrage de Hunston & Francis (2000), et qui s'inscrit résolument dans une perspective lexicologique⁴⁶. Dans le droit fil des travaux de Sinclair, elle se propose en effet de repérer puis d'étudier les divers patrons distributionnels des mots, un *pattern* n'étant, dans cette optique, qu'une combinaison syntaxique dans laquelle peuvent figurer des classes sémantiquement homogènes de mots. Pour Hunston & Francis (2000: 37):

The patterns of a word can be defined as all the words and structures which are regularly associated with the word, and which contribute to its meaning. A pattern can be identified if a combination of words occurs relatively frequently, if it is dependent on a particular word choice, and if there is a clear meaning associated with it:

Le recensement des patterns est, répétons-le, la recherche systématique des patrons distributionnels des mots comme des adjectifs, des adverbes, des noms, etc. Il s'agit d'identifier l'environnement syntaxique d'un item lexical donné – ses *patterns grammaticaux* - et pour ce faire, on procède en deux étapes. La première étape consiste à repérer l'association fréquente entre un *mot cible* et d'autres mots (les *collocations* de ce mot) dans des structures pertinentes pour l'interprétation de ce mot cible (ses *colligations*). La seconde étape est l'identification proprement dite des patterns lexicaux

⁴⁵ Legallois & François (2006 : 28) proposent le terme de *grammaire contextualiste* pour cette approche.

⁴⁶ On trouvera un exposé très clair et documenté des grammaires de patterns dans Legallois & François (2006 : 28-48). Sur l'histoire du contextualisme britannique, cf. William (2003).

récurrents, sachant qu'un même mot peut posséder plusieurs patterns, et qu'un même pattern grammatical peut être associé à plusieurs mots (*op. cit.*: 37):

The procedure for investigating the patterns of a word involves selecting at random a number of concordance lines and sorting them into alphabetical order. They may be right-sorted or left-sorted according to the word one is examining. In the case of a verb, it is more revealing to sort to the right, since most verbs have complementation patterns which follow them (*he decided to leave, he hated leaving*). Though complementation patterns are usually the most interesting facts about verbs, there may be reasons for sorting them to the left, as this would show how often a verb occurs in the passive or infinitive, which modals it is often used with, what are its typical Subjects, whether it is frequently negative and so on. In the case of nouns, too, sorting to the right reveals their complementation.

Par ailleurs (*op. cit.*: 43):

Just as a word can have several different patterns, so a pattern can be seen to be associated with a variety of different words.

Résumons : les patterns grammaticaux sont établis à partir des classes grammaticales des mots, et jamais à partir de leur(s) fonction(s). De telles classes grammaticales correspondent donc à une analyse strictement linéaire, non-fonctionnelle, et en particulier non hiérarchique. Par ailleurs, la grammaire des patterns est une théorie lexicale dans la mesure où elle décrit un mot par un ensemble de tels patterns grammaticaux récurrents. Là s'arrête la ressemblance avec les phrases minimales de M. Gross ou encore les fonctions lexicales de Mel'čuk, puisqu'il n'est envisagé aucun passage d'un pattern à un autre sur le principe par exemple d'une transformation, comme ce serait le cas dans des grammaires d'inspiration générativiste.

Voici maintenant un exemple de recherche de pattern, celui de la construction V n n (*i.e.* Verbe + Nom + Nom)⁴⁷, que nous réduirons à l'essentiel, et qui apparaît dans trois structures (exemples de Hunston & Francis, 2000 : 87-90) : *I gave him some bread* (verbes 'ditransitifs'), *She called them all idiots* (V + Objet + Cplt d'objet), et *They beat us three-nil* (V + Objet + Quantité). Or les verbes appartenant à ce pattern présentent des caractéristiques sémantiques spécifiques selon la structure envisagée. Sans entrer dans le détail, les verbes dits 'ditransitifs' sont principalement des verbes qui : a) concernent le fait de donner ou de refuser quelque chose à quelqu'un (*to promise, to refuse*) ; b) concernent le fait de faire quelque chose pour quelqu'un (*to book, to sing*) ; c) concernent le fait de communiquer quelque chose à quelqu'un (*to teach, to write*) ; d) concernent le fait de procurer à quelqu'un un avantage ou un désavantage (*to charge, to spare*) ; e) concernent des sentiments et des attitudes (*to envy, to forgive*). Les verbes suivis de (Objet + Cplt d'objet) sont essentiellement des verbes concernant le fait de classer quelque chose dans une catégorie donnée, soit par dénomination, soit par attribution (*to name, to consider*) ; de situer quelqu'un ou quelque chose dans une position particulière (*to anoint*) ; de situer quelqu'un ou quelque chose dans une position particulière par opinion (*to adjudge*) ou par un processus causal (*to make*). Enfin, les verbes de la classe

⁴⁷ Dans la notation de Hunston & Francis (*op. cit.* : 33), un pattern est représenté comme une séquence d'éléments. Le mot cible est en majuscule, les autres éléments en minuscules.

Verbe + Objet + Quantité sont exclusivement des verbes concernant le fait de gagner ou de perdre dans un jeu ou dans un sport (*to beat, to lose, to win*). On constate dans cet exemple que le résultat du recensement des patterns est conforme aux objectifs de la théorie proposée : les patterns mis en évidence révèlent une connexion entre une structure (ou une sous-structure) syntaxique et une certaine signification générale. De ce point de vue, les tenants de la *pattern grammar* mettent à mal l'hypothèse générativiste d'une séparation totale entre syntaxe et sémantique. D'autant plus que le travail sur grands corpus évite de postuler des hypothèses hasardeuses à partir d'exemples inventés sans vérification des fondements de ces exemples. Enfin, la méthode des *pattern grammars* replace la phraséologie au centre de la langue, en accord avec l'*idiom principle* de Sinclair. Les éléments phraséologiques ne sont pas des phénomènes marginaux et rares, ils sont au contraire fréquents, et constituent donc un des modes de production de la langue.

On peut cependant adresser un certain nombre de critiques à la position de la *pattern grammar*. La première est que l'usage de grands corpus ne met pas à l'abri des hypothèses – un corpus est et reste un construit, et encore moins si ces corpus sont écrits et véhiculent alors certaines thèses des grammaires normatives. Une seconde critique est le caractère très flou des gloses utilisées pour décrire les patterns dégagés. Comment sait-on par exemple, que tel verbe signifie 'classer un objet dans une catégorie X' autrement que par une espèce de connaissance intuitive de la langue, i.e. sans partir de définitions des gloses proposées et des critères afférents? Tout le problème de la *pattern grammar* est qu'elle estime que sa façon de procéder la met à l'abri de toute approche théorique. Et tout ce que cette approche peut dire, c'est que, contrairement aux grammaires de construction selon Goldberg, la grammaire de pattern n'a nul besoin d'une justification *externe* de sa validité, à savoir une certaine *réalité cognitive*, et s'appuie sur une seule réalité empirique et statistique, qui est en fait considérée comme *interne*. Pour la grammaire de pattern, la vérité, c'est l'usage : en fait, ce qui est dénommé *usage* est une extrapolation à partir d'un nombre fini d'exemples, quelque grand qu'il soit. Dernier point enfin, valable également pour les approches constructionnistes : aucune des deux approches ne semble pouvoir expliquer pourquoi il y a des constructions idiomatiques, quel que soit le nom qu'on leur donne, en particulier pourquoi certaines d'entre elles sont productives, ce qui semble, en partie du moins, contraire à la notion même de figement.

5.3. Matrices lexicales et sous-systèmes lexicaux

Les notions de *matrice lexicale* et de *sous-système lexical* sont nées de l'étude complète – y compris diachronique – de certaines formations lexicales particulières, à savoir les composés nominaux de type N_1 à N_2 (Anscombre, 1990, 1999, 2011), les proverbes (2011, 2019a) et les expressions du type *faire ses adieux* (2022b, 2022c) pour le français, et enfin les noms de bruit en *-ido* pour l'espagnol (Anscombre, 2024). Ces formations présentent un certain nombre de caractéristiques qui rappellent les problèmes de figement et d'idiomaticité, d'où l'intérêt d'une étude complète et détaillée de ces sous-systèmes lexicaux. Nous parlons de *sous-systèmes lexicaux*, car nous nous proposons de montrer que l'avènement de leur formation suit des règles tout à fait spécifiques, et qui ne sont pas sans rappeler sur certains points les constructions de Goldberg.

5.3.1. Des proverbes aux composés N_1 à N_2

Les N_1 à N_2 sont une série de noms complexes formés à partir de deux autres noms, selon un procédé morphologique qui consiste à les relier par la préposition *à*. Dans la plupart des cas, le second substantif n'est précédé d'aucun déterminant, les exceptions étant rares : *boîte aux lettres*, *foire aux livres*, *soupe aux légumes*, etc. Ces composés nominaux, été classés par Benveniste (1966 : 163-171) dans les phénomènes de *synapsie*, possèdent deux propriétés :

- a) Il s'agit d'une formation typiquement française à l'intérieur des langues romanes. Bien que possible par exemple en espagnol, elle est très rare, et ce n'est que récemment que certains termes techniques ayant cette forme sont apparus, à l'instar – voire inspirées – du français : ainsi *avión a reacción* (= avion à réaction), *dibujo a lápiz* (= dessin au crayon), *navegación a vela* (= navigation à voile), *taladro a acero rápido* (= mèche à acier rapide), etc. ;
- b) L'essentiel de ces termes concerne des objets, en particulier des objets techniques : *imprimante à jet d'encre*, *stylo à bille*, *flûte à champagne*, *boîte à outils*, *immeuble à grand standing*, etc. Il y a peu de composés [+Humain] ou même [+Animé], environ une trentaine : *homme à femmes*, *fille à soldats*, *auteur à succès*, *fils à papa*, etc. De plus, cette sous-série est très peu productive, et comprend beaucoup de termes métaphoriques probablement en voie de figement. Ce qui confirme l'idée, fréquemment avancée, d'une relation entre le figement et la non-productivité.

Une caractéristique de la sous-catégorie [-Humain] de tels composés est qu'elle est en revanche extrêmement productive en français contemporain, ce qui est contraire aux hypothèses de figement qu'on a parfois proposées pour expliquer leur formation, ainsi G. Gross (1988, 2012). En fait, cette hypothèse d'un plus ou moins grand figement de nos noms composés repose sur la constatation d'une certaine rigidité syntaxique et sémantique. Bien entendu, comme toute catégorie linguistique, celle de nos composés exhibe des éléments plus ou moins figés – par exemple *sac à vin* ou *pot à tabac*, mais G. Gross n'envisage pas l'hypothèse – qui est celle des constructionnistes – d'une catégorie lexicale où forme et sens sont liés et se contraignent réciproquement, et sont responsables des phénomènes signalés.

Pour rendre compte de ces phénomènes, nous aurons recours à la notion de *matrice lexicale*, présentée dans Anscombe (1990), et développée par la suite en particulier dans Anscombe (2011, 2019a). Nous appellerons *matrice lexicale* un schéma tel que :

- (a) Il comporte des unités linguistiques fixes et des variables linguistiques
- (b) Les unités linguistiques fixes sont très généralement des éléments grammaticaux.
- (c) Les variables linguistiques représentent des éléments lexicaux.
- (d) Ce schéma est productif, du moins dans une certaine mesure.
- (e) À ce schéma est attaché un invariant sémantique reliant les unités lexicales en présence, qui n'est pas déductible des seuls termes en présence et impose entre autres des relations préexistantes entre les unités lexicales. Il y a donc un lien entre forme et sens.

Nous illustrerons brièvement cette notion avec un exemple que nous avons beaucoup étudié, celui des proverbes. Plus particulièrement, nous nous centrerons sur les proverbes du type *Qui va à la chasse perd sa place*. On remarque immédiatement que ce schéma est très bien représenté dans le monde parémique français, et productif : *Qui casse les pots les paie* ; *Qui dort dîne* ; *Qui ne dit mot consent* ; *Qui n'entend qu'une cloche n'entend qu'un son* ; *Qui ne risque rien, n'a rien* ; *Qui veut la fin, veut les moyens* ; etc. Il est tentant d'y voir une matrice lexicale de type $Qui\ GV_1,\ GV_2$, le *qui* étant l'élément (grammatical) fixe, et les GV_1 et GV_2 les éléments lexicaux variables. Quant à l'invariant sémantique, il sera que GV_1 et GV_2 sont liés par une relation de généricté : 'c'est toujours le cas que si GV_1 , alors GV_2 ' (sans entrer dans le détail, rappelons pour mémoire que ce *qui* signifiait à l'origine 'si on'). A chaque couple (GV_1 , GV_2) spécifique correspondra un proverbe, et tant que et la fixité de la forme et la relation de généricté seront préservées, d'éventuelles variations seront possibles, ainsi : *Qui va à la chasse, peut perdre sa place* ; *C'est celui qui casse les pots qui doit les payer* ; *Qui ne risque rien n'obtient rien* ; *Qui veut la fin, met les moyens* ; etc. (variantes effectivement attestées, parmi d'autres) On comprend alors la possibilité de variantes, plus ou moins nombreuses selon les cas. Mais également l'existence de contraintes liées à la coexistence dans la matrice d'un invariant sémantique et d'une structure partiellement fixe. On remarque également qu'à chaque schéma va correspondre un ensemble de proverbes 'apparentés', qui forment en quelque sorte un sous-système lexical. L'ensemble des proverbes est donc l'ensemble de ces sous-systèmes: (cf. chez Goldberg (2006-16) la notion de *family of constructions*). Gómez-Jordana (2012) a recensé pour le français et l'espagnol une douzaine environ de matrices lexicales, certaines plus productives que d'autres. Outre que le français tend à privilégier les proverbes abstraits, certaines structures semblent en voie de disparition, ainsi *Tel x, tel y*, limitée aujourd'hui à trois éléments : *Tel est pris qui croyait prendre* ; *Tel père, tel fils* ; *Tel qui rit vendredi, dimanche pleurera*. On voit à la lumière de ces faits une possible définition du figement : une expression figée est une expression qui appartient à une sous-structure lexicale qui n'est plus productive. En revanche, une expression qui appartient à une sous-structure lexicale productive n'est pas une expression figée, et peut très bien comporter des spécificités (comme la non-compositionnalité) dues au lien étroit que présentent sa forme et son sens. Une des questions que pose la notion de matrice lexicale sont les raisons qui déclenchent la 'fabrication' de telles matrices. Si on peut raisonnablement penser, dans le cas des N_1 à N_2 , à une nécessité liée à l'apparition d'une ou plusieurs technologies – cette composition est très productive dans de tels domaines, le motif qui pousse à créer de nouveaux proverbes semble plus obscur, et a peut-être à voir avec les codes d'appartenance à telle ou telle autre communauté linguistique.

Dans l'optique théorique des matrices lexicales, on voit le traitement qui peut être proposé pour les N_1 à N_2 : il s'agira de les considérer comme issus de matrices lexicales. La base sera une forme à élément fixe – la préposition *à* – et à éléments (lexicaux) variables N_1 et N_2 . Quant à l'invariant sémantique attaché à cette forme, il postule qu'un N_2 ne peut figurer dans un schéma N_1 à N_2 que si certains N_1 présentent avec le N_2 envisagé un lien constitutif⁴⁸. Un exemple fera mieux comprendre ce dont il s'agit : dans

⁴⁸ Ce lien est ce que nous avons appelé ailleurs (Anscombe, 1999) une *propriété intrinsèque* (i.e. constitutive) *accidentelle* (tous les N_1 ne la possèdent pas).

une chemise à manches courtes, les manches courtes sont présentées comme définissant, à l'intérieur de la classe des chemises, une sous-classe dont la caractéristique définitoire sont les manches courtes. Et une telle contrainte doit être satisfaite pour autoriser le composé : *une chemise à manches déchirées* est difficile – on dirait *une chemise aux manches déchirées* – sauf dans un contexte où le dernier cri de la mode serait d'avoir les manches de chemise déchirées, auquel cas ce trait deviendrait caractéristique d'un certain type de chemise, donc définitoire, et le composé serait alors viable. Dernier point déjà mentionné et sur lequel nous n'insisterons pas : dans un composé $N_1 à N_2$, les possibilités de reprise de N_2 par anaphore associative révèlent le type de relation qui unit N_1 et N_2 , et qui n'est pas nécessairement visible directement en structure de surface : il peut s'agir d'une relation locative (*réservoir à essence*), stative (*stylo à bille*) et enfin processive (*type à histoires*). Cette relation constraint le composé considéré au niveau des propriétés syntaxiques et sémantiques, sans qu'il soit pour autant nécessairement figé : comme indiqué *supra*, la plupart des composés de ce type font partie de séries productives. Enfin, puisqu'il y a trois types de relation possibles entre N_1 et N_2 , l'ensemble des composés de ce type est donc une *famille lexicale* composé de trois sous-structures lexicales.

5.3.2. Les expressions idiomatiques du type *faire ses adieux*

On classe habituellement parmi les *pluralia tantum* une sous-classe dont l'essentiel des éléments figure dans la liste suivante : *jurer ses grands dieux*, *faire ses adieux*, *présenter ses hommages*, *présenter ses excuses*, (*adresser*) *ses félicitations*, (*présenter*) *ses meilleurs vœux*, *présenter ses condoléances*, *adresser ses remerciements*, (*adresser*) *tous ses regrets*, (*adresser*) *ses sentiments les plus respectueux*, *adresser ses compliments*, *faire ses amitiés à quelqu'un*, (*adresser*) *ses salutations les meilleures*, *arriver à ses fins*, *assumer ses responsabilités*, *faire ses confidences*, *faire ses débuts*, *faire ses premiers pas*, (*prendre + assumer*) *ses responsabilités*, *jurer ses grands dieux*, *reprendre ses droits*, etc.

Les éléments de cette liste se caractérisent par les propriétés suivantes : (i) Le groupe nominal (au 'pluriel') n'est pas un nom collectif au sens de 'ensemble d'éléments discrets' ; (ii) Ce groupe nominal est généralement combiné avec des verbes 'dédiés', en très petit nombre, parfois un seul, dont beaucoup de verbes-supports ; (iii) Ce groupe nominal comprend un possessif, qui renvoie obligatoirement au sujet du verbe ; (iv) Ce groupe nominal a une signification particulière, qui n'est pas celle de la version au singulier quand cette dernière existe. En particulier, la version au singulier n'est pas combinable avec les verbes dédiés ; (v) Même si les deux versions singulier et pluriel existent en synchronie, la version 'plurielle' est diachroniquement toujours postérieure à la version 'au singulier'.

Ces caractéristiques posent la question de savoir s'il s'agit ou non de *pluralia tantum*. Pour Anscombe (2009, 2022b, 2022c), il s'agit en fait d'une extension du *-s* dit adverbial, ou plus exactement, du recyclage du *-s adverbial* comme marque d'un schéma lexical destiné à former les éléments ci-dessus, des matrices lexicales donc. Nous rappellerons brièvement l'odyssée du *s-adverbial* ; il s'agit au départ d'une fonction adverbiale, productive dès l'ancien français, et servant à former des adverbes, ainsi

meismes, volentiers, par amors, etc., dont certains ont subsisté jusqu'à aujourd'hui (*alors, toujours, certes, etc.*). On retrouve ce phénomène dans beaucoup de langues indo-européennes, en particulier les langues romanes. Il a longtemps été considéré comme relevant d'états antérieurs de la langue, mais a été recyclé dans plusieurs langues contemporaines – dont le français, l'espagnol, l'anglais, pour former différentes unités lexicales : interjections, locutions adverbiales dérivé d'interjections, surnoms, adjectifs marqués comme *craignos* en français, etc. C'est l'idée adaptée de Lass (1990) qu'on est passé d'un *-s* comme marque flexionnelle issue du latin à un *-s* dérivationnel, puis à une marque de lexicalisation (Anscombe 2022b), apte donc à servir dans la création de sous-systèmes lexicaux, en particulier sur la base de matrices lexicales. Dans le cas qui nous occupe, le processus de formation serait *grosso modo* le suivant :

- a) Dans un premier stade, existe une combinaison de type Vsup + possessif + GN (sing./plur.).
- b) Dans un domaine particulier, GN tend à prendre de façon systématique un sens spécifique qui n'est pas le sens usuel. Par exemple *début* 'première fois dans le monde ou au théâtre' face à *début* 'première réalisation d'une action', ou encore *excuse* 'action de s'excuser' face à *excuse* 'explication de la cause'.
- c) La version au pluriel tend à s'imposer sur la version au singulier pour les sens spécifique, puis à l'éliminer.
- d) Cette version plurielle perd son sens collectif, et en vient à s'aligner sur le paradigme des locutions de type V + possessif + GN-*s*, où le singulier n'est pas possible – du moins avec le sens spécifique , et où V est peu susceptible de variations.

Voici à titre d'exemple, et réduit à ses grandes lignes, le parcours qui conduit de *début* à faire ses débuts tel qu'il apparaît au travers des éditions successives du *Dictionnaire de l'Académie française* (s.v. *début*) :

- Editions de 1693, 1718, 1740, 1762: « Le premier coup à certains jeux, comme au mail, à la boule &c. *Voilà un beau début, faire un beau debut* ».
- Edition de 1798 : « Le premier coup à certains jeux, comme au mail, au billard, à la boule, etc. *Voilà un beau début. Faire un beau début* [...] Il se dit particulièremt. Des premiers temps employés par les Comédiens pour essayer leurs talens sur l'un des théâtres de la Capitale [...] *Son début a duré trois mois* ».
- Edition de 1835, 1878 : « Le premier coup à certains jeux, comme au billard, à la boule, au mail, etc. *Voilà un beau début. Faire un beau début* [...] Il se dit également de La manière dont on commence un genre de vie, une entreprise ; des premiers actes qu'on fait dans un emploi, dans une profession. *Son début dans le monde ne dut pas prévenir en sa faveur* [...] Il se dit, particulièrement, De ceux qui paraissent pour la première fois sur le théâtre, et Des acteurs qui essayent leurs talents sur un théâtre ou ils n'avaient point encore paru. *Un rôle de début. Premier, second début. Il a obtenu un ordre de début. Cet acteur a terminé ses débuts. Ses débuts ont duré trois semaines. Faire son début, ses débuts* ».

- Edition de 1932-35 : « Commencement d'une chose. *Au début de la maladie. Le début d'un discours. Un début qui promet. Un incident se produisit au début de la séance.* Il se dit spécialement à propos des Premiers essais sur le théâtre ou sur un théâtre. *Un rôle de début. Premier, second début. Il a obtenu un ordre de début. Cet acteur a terminé ses débuts. Ses débuts ont duré trois semaines [...] Faire son début dans le monde* ».
- Edition de 1932-35 Souvent au pluriel. Action de débutter ; première étape d'une carrière, surtout en parlant d'un artiste. *De bons, de mauvais débuts. Ses débuts au théâtre sont des plus prometteurs. Un début dans la vie. Il n'en est pas à son début.*
- Edition en ligne (9^{ème} édition) : [...] Action de débutter ; première étape d'une carrière, surtout en parlant d'un artiste. *De bons, de mauvais débuts. Ses débuts au théâtre sont des plus prometteurs. Un début dans la vie. Il n'en est pas à son début. Faire ses débuts dans les lettres, dans la politique, dans les affaires, etc. Cette actrice fit ses débuts en tenant le rôle d'Agnès dans « L'École des femmes ». Par analogie. Faire ses débuts dans le monde, faire ses premiers pas dans la vie mondaine, aller à son premier bal.*

Ce type de phénomène est typique de la formation de matrices lexicales ou de constructions selon Goldberg. Quand s'est formée cette sous-catégorie lexicale ? Au départ, un exemple isolé, celui de *reprendre ses droits*, très vraisemblablement. Originellement, *droit* est essentiellement singulier, et apparaît au pluriel uniquement pour signifier les droits du roi ou de la couronne. Vers 1457, on trouve la combinaison *avoir ses droits* prédiquée de la nature. Cette expression aura servi de modèle pour un schéma lexical de type Vsup + possessif + GN-s, servant à former des matrices lexicales dans un domaine particulier, celui des relations sociales en général, comme on peut le voir sur la liste *supra*. Le -s final du GN sert finalement à signaler que ces matrices lexicales concernent la vie sociale en général, soit relationnelle (*faire ses excuses, présenter ses hommages*), soit mondaine (*faire ses débuts dans le monde*), soit enfin juridique (*assumer ses responsabilités*). Il atteste donc de la présence d'une sous-structure lexicale, faite d'une famille de matrices où sens et forme sont liés d'une façon spécifique.

5.3.3. Les noms de bruits en *-ido* en espagnol

Ce cas est intéressant à plus d'un titre, en particulier le fait qu'il s'agit d'un cas curieux de morphologie synthétique, alors que la majorité des cas d'idiomaticité traités relèvent de la morphologie cette fois analytique, ainsi les locutions verbales comme celles que nous venons d'examiner⁴⁹. Par ailleurs, la langue concernée est l'espagnol et non pas le

⁴⁹ Un cas de création suffixale nous paraît relever d'un mécanisme analogue : celui des adverbes en *-ment* en français, sur la base de la description très détaillée qui est donnée de ce phénomène dans la *Grande Grammaire Historique du Français* (2020), Christiane Marchello-Nizia, Bernard Combettes, Sophie Prévost & Tobias Scheer (éds.), De Gruyter, Mouton : § 34.4.2.

français, ce qui montre que notre analyse reste valide au-delà de la simple considération du français

De quoi s'agit-il ? L'espagnol possède une liste de noms d'action qui désignent tous des noms de bruits, de cris d'animaux ou de cris humains, qui est la suivante: *aullido*, *balido*, *berrido*, *bramido*, *bufido*, *chasquido*, *chillido*, *chirrido*, *crujido*, *estallido*, *gamitido*, *gañido*, *garlido*, *gemido*, *graznido*, *gruñido*, *hipido*, *ladrido*, *latido*, *maullido*, *mugido*, *pitido*, *piulido*, *plañido*, *quejido*, *rechinido*, *relinchido*, *resoplido*, *ronquido*, *silbido*, *sonido*, *soplido*, *susurrido*, *tañido*, *traquido*, *tronido*, *zumbido*, *zurrido*. Par exemple, *balido* 'bêlement', *pitido* 'sifflement', *gemido* 'gémissement', *crujido* 'cristissement', *maullido* 'miaulement', *ronquido* 'ronflement', *zumbido* 'vrombissement'. Leurs caractéristiques sont les suivantes :

- a) Ils représentent exclusivement des cris d'animaux ou d'humains, des sons inarticulés, des bruits produits par des objets, comme étant des habitualités dispositionnelles.
- b) En tant que noms de bruits, ces substantifs admettent la combinaison avec des verbes 'dédier' correspondant pour l'espagnol aux deux notions générales 'émettre' ou 'capter'. Entre autres, *dar* 'donner', *emitir* 'émettre', *lanzar* 'lancer', *sonar* 'retentir' (en émission) ; *escuchar(se)* 'percevoir', *oír(se)* 'entendre' (en réception).
- c) Ils sont formés par suffixation en *-ido* sur une base qui fournit par ailleurs des verbes de sens apparenté. Ainsi *aull-ido/aull-ar*, *bal-ido/bal-ar*, *la--tido/lat-ir*, *sopl-ido/sopl-ar*, etc.
- d) Ce suffixe *-ido* est le même quelle que soit la terminaison (*-ar*, *-er* ou *-ir*) du verbe apparenté formé sur la même base. Il ne peut, donc s'agir d'une nominalisation d'un participe passé, puisque la voyelle accentuée de ce dernier varie avec la terminaison verbale. Alors que *estall-ar* 'éclater', a pour participe passé *estall-ado* 'éclaté', le nom de bruit correspondant est *estall-ido* 'éclatement'.
- e) Après avoir été très productive, cette dérivation semble ne plus l'être en espagnol contemporain.

Sans entrer dans le détail (cf. Craddock & Georges, 1963-64), les auteurs ayant étudié ce phénomène y voient une formation par analogie avec le latin, qui formait des noms de bruit en *-ītūs*, *-ūs* ou *-ītūs*, *-ūs*, à partir essentiellement de verbes de la quatrième conjugaison (celle de *audio* 'entendre'), ainsi *aud-io/aud-ītūs* 'action d'entendre' .

Nous proposons de voir dans l'apparition et l'installation de cette catégorie la création d'une sous-structure lexicale qu'on peut représenter par le schéma lexical *Base verbale + -ido*, *Base verbale* signifiant 'base sur laquelle sont également formés des verbes', avec comme invariant sémantique 'nom de bruit correspondant au verbe apparenté de même base verbale, et désignant une disposition spécifique'. Le processus à l'œuvre serait le suivant : au départ (début XIII^e siècle), apparaissent des noms de bruit sans verbe apparenté, probablement par imitation des noms de bruit du latin en *-ītūs*, *-ūs* ou *-ītūs*,

*-ūs*⁵⁰ : *alarido* ' hurlement' et *ruido* ' bruit'. Puis sont forgés, vers la fin du XIII^e siècle, *bramido* ' brame', *ladrido* ' aboiement', *sonido* ' son' et enfin *tronido* ' grondement (de tonnerre)' (début de la seconde moitié du XIII^e s.), avec cette fois les verbes correspondants *bramar*, *ladrar*, *sonar* et *tronar*. La langue tend donc à regrouper des cas isolés présentant une certaine analogie en les intégrant dans une même structure 'préfabriquée'. Il s'agit en quelque sorte d'une manifestation par la langue d'autorégulation. Une seconde période, qui va durer jusqu'à l'aube du XVII^e siècle, est la période de productivité, qui voit la production de noms de bruit exclusivement par application du suffixe *-ido* à des verbes majoritairement en *-ar*⁵¹. À partir du XVII^e siècle, la productivité devient faible, pour ne pas dire inexistante, ce qui est peut-être dû au fait qu'il s'agit d'un lexique très spécialisé mais non technique, et donc peu sollicité dans la vie courante. Quoi qu'il en soit, ce parcours illustre de façon idéale ce que sont selon nous l'*idiomaticité* et le *figement*. L'*idiomaticité* est la création d'une sous-structure lexicale par regroupement de cas isolés, regroupement qui se fait par des schémas 'préfabriqués' (cf. Gülich & Kraft, 1997, pour un des premiers textes en faveur de cette position) à forte cohésion sens/forme, souvent repérables par des marques spécifiques. Les structures figées sont ce qui reste de ces *préfabs* lorsqu'ils cessent d'être productifs.

6. CONCLUSIONS(S) ET PROSPECTIVE

À la suite de ce long exposé, il nous semble qu'un certain nombre de conclusions s'imposent concernant la nature de ce qu'il est convenu d'appeler le *figement* ou encore l'*idiomaticité*, bien que, on l'aura compris, ces deux notions ne coïncident pour nous que partiellement. En fait, tout vient de ce qu'on a considéré dès le départ que les *expressions figées* s'opposaient aux *expressions libres* en ce que seules les secondes étaient normales, les premières étant alors, par voie de conséquence, hors système linguistique, au moins partiellement. On ne pouvait donc au mieux que les recenser et les classer selon leurs particularités. Or, une telle opposition n'est pas un observable, mais le résultat d'hypothèses générales sur le modèle choisi pour décrire une langue en général et de tels phénomènes en particulier. En effet, les expressions figées présentent des anomalies et/ou des contraintes au niveau syntaxique et au niveau sémantique. Un modèle linguistique où syntaxe et sémantique sont deux modules indépendants ne peut donc rendre compte de ces expressions figées qu'en les considérant comme au moins partiellement hors-système. C'est le cas de la grammaire générative transformationnelle, où la syntaxe est totalement indépendante de la sémantique. C'était déjà en grande partie la position de nos grammaires scolaires et universitaires, qui nous enseignent la maîtrise de la syntaxe, étant bien entendu qu'une fois acquise une telle maîtrise, celle de la sémantique s'ensuit. Pour les amateurs de constructions, collocations et autres matrices lexicales, syntaxe et sémantique sont deux modules interdépendants. Rien n'empêche alors de considérer la

⁵⁰ Comme le révèle l'apparition vers 1260 de *gémito* 'gémissement', formé à partir du latin *gēmītūs*, *-ūs*, accentué sur l'antépénultième comme son ancêtre latin, alors que l'actuel descendant *gemido* de même sens est accentué cette fois sur la pénultième. Pour les détails, cf. Anscombe (2022c).

⁵¹ Les verbes en *-ar* représentent en espagnol la conjugaison la plus régulière et aussi de très loin la plus productive.

formation de certaines unités lexicales comme alliant de façon indissociable syntaxe et sémantique. De ce point de vue, l'idiomaticité n'est pas un phénomène syntaxique, mais à la fois syntaxique, sémantique et lexical. Elle est d'abord sémantique (les *collocations*), et c'est le lien entre syntaxe et sémantique qui provoque en retour des contraintes au niveau syntaxique. Dans le cas par exemple des matrices lexicales, c'est l'invariant lexical de la matrice qui contrôle les variations d'une forme qui renvoie à cet invariant. Enfin, la notion de *figement* est d'une certaine façon un faux-problème. L'option qu'ont choisi les tenants des constructions, patterns et autres matrices lexicales, est une option à la Sinclair. Ils considèrent que de tels schémas lexicaux font partie intégrante de la langue, au même titre que les constructions dites libres. Leurs spécificités est la production de lexiques spécialisés, c'est-à-dire relatifs à un domaine particulier. On notera qu'il n'y a d'ailleurs là rien d'étonnant : ainsi, le lexique technique actuel du français forme de façon ultra-majoritaire des noms d'action par suffixation en *-age*, ce qui n'a pas toujours été le cas. D'où deux remarques finales : ces constructions 'fixes' signalent de par leur forme leur domaine d'application, et les critères signalant la singularité de tels schémas lexicaux sont propres à chaque sous-catégorie à laquelle ils renvoient. Ce qui explique qu'on n'ait jamais pu trouver de critères 'généraux'. D'autre part, le figement n'est pas un phénomène linguistique central : il est le résultat de l'existence de tels schémas lexicaux lorsqu'ils cessent d'être productifs. Le fonctionnement, la productivité et les propriétés de tels schémas relèvent de la synchronie, leur existence et la perte de productivité qui conduit à un figement cette fois de la diachronie.

Il nous semble donc que les études sur le figement doivent opérer, de ce point de vue, un changement radical. D'une part, un retour sur les fondements de la notion même de figement nous semble s'imposer, nous pensons l'avoir montré au moins partiellement. D'autre part, la majorité des études sur les expressions dites figées concernent les locutions verbales et, dans une moindre mesure, les locutions adverbiales. Enfin, la comparaison entre les usages libres et les usages dits figés semble avoir été négligée, alors qu'elle pourrait s'avérer décisive pour la détermination de l'appartenance à telle ou telle autre sous-catégorie lexicale. Deux domaines nous paraissent avoir été particulièrement négligés : d'une part, celui des créations monolexicales, comme les noms de bruit en espagnol, mentionnés dans ce texte, qui ne semblent avoir intéressé jusqu'à présent que les philologues. Et dans un domaine dernièrement très à la mode, celui des marqueurs de discours, où la vulgate - nous sommes tous coupables - semble implicitement considérer que la formation d'un nouveau marqueur de discours correspond à un cas de figement. Nous avons-nous-même conclu à un figement en cours de la formule française *je me suis laissé dire* (Anscombe 2023, sous prétexte qu'elle présentait de faibles possibilités de variation). Il nous semble aujourd'hui qu'un autre angle d'attaque devrait être adopté. Dans le même ordre d'idée, la création d'un nouveau marqueur par mise en incise est souvent considérée comme accompagnée d'un figement, sans qu'on ait approfondi ces affirmations. C'est entre autres le cas des marqueurs formés sur des verbes dits *recteurs faibles*, ainsi *croire* ou *penser* (Blanche-Benveniste & Willem, 2016), sans qu'on se soit jamais interrogé sur l'origine et le statut de leur productivité dans le domaine des marqueurs. Citons également pour mémoire des séries comme (*Bon alors + Ben alors + Toi alors + Et alors + Ça alors + Zut alors + ...*), etc.

Le chapitre du figement et de son statut est loin d'être clos, et a encore de beaux jours à vivre.

Jean-Claude Anscombe
Directeur émérite (LT2D)
U. de Cergy-Pontoise)

BIBLIOGRAPHIE⁵²

- Alster, B. (1997), *The Wisdom of Ancient Sumer*, CDL Press, Bethesda, Maryland.
- Anscombe, J.C. (1981), "Marqueurs et hypermarqueurs de dérivation illocutoire : notions et problèmes ", *Cahiers de linguistique française*, n° 3, 75-123.
- Anscombe, J.C. (1986), "L'article zéro en français : un imparfait du substantif ?", *Langue française*, 72, 4-39.
- Anscombe, J.C. (1990), "Pourquoi un moulin à vent n'est pas un ventilateur", *Langue française*, n°86, 103-125. Anscombe, J.C. (1999), "Le jeu de la prédication dans certains composés nominaux", *Langue française*, n°122, 52-69.
- Anscombe, J.C. (2001), "Dénomination, sens et référence dans une théorie des stéréotypes nominaux", in *Linguistique de la dénomination*, *Cahiers de praxématique*, n° 36, 43-72.
- Anscombe, J.C. (2008), "Les comparatives du type *être Adj. comme P* : des tournures figées ou non ?", in *Les séquences figées : entre langue et discours*, Peter Blumenthal & Salah Mejri (éds.), *Zeitschrift für Französische Sprache und Literatur*, Cahier n°36 : 13-25.
- Anscombe, J.C. (2011), "Figement, idiomatique et matrices lexicales", in *Le figement linguistique : la parole entravée*, Jean-Claude Anscombe & Salah Mejri (éds.), Honoré Champion Editeur, Paris : 17-40.
- Anscombe, J.C. (2012), "Pour une théorie linguistique des phénomènes parémiques", *La parole exemplaire*, Jean-Claude Anscombe, Bernard Darbord & Alexandra Oddo (éds.), Coll. *Recherches*, Armand Colin, 21-39.
- Anscombe, J.C. (2019a), "Figement, lexique et matrices lexicales", *Cahiers de lexicologie*, 1, n° 114, *Les phrases préfabriquées : sens, fonctions, usages*, Gaétane Dostie & Agnès Tutin (dirs.) : 119-147.
- Anscombe, J.C. (2019b), "Proverbes : si les vulgates m'étaient contées", *Actes du congrès de Grenoble, Proverbes et locutions figées. Description et catégorisation* (Salam Diab-Duranton, Georges Kleiber & Abdendi Lachkar (éds.), Guenthner, Paris : 17-66.

⁵² Cette bibliographie n'a aucune prétention à l'exhaustivité. Elle signale les textes soit fondateurs d'une ou plusieurs nouvelles orientations, soit pédagogiquement incontournables (ainsi Legallois & François 2006), soit enfin porteurs d'exemples ou de définitions significatifs pour notre propos – ainsi certains dictionnaires.

- Anscombe, J.C. (2020), " La marque d'intensification des comparatives proverbiales en français et en espagnol. Etude et genèse des différences et des ressemblances", *Linguisticae Investigationes : Les variations diatopiques dans les expressions figées*, Pedro Mogorron, Aude Grezka & Lucía Navarro-Brotóns (dirs.), 43, 2 : 172-189.
- Anscombe, J.C. (2022a), "Usage et usages en parémiologie linguistique ", in *Les usages de l'usage*, Béatrice Wendling & Sandra Laugier (dirs.), Série *Les concepts fondateurs de la philosophie du langage*, vol. 9, ISTE Editions, London : 221-238.
- Anscombe, J.C. (2022b), "La nature reprend ses droits : un pluriel bien singulier", *Cahiers de lexicologie*, 1, n° 120, *Varia* : 11-36.
- Anscombe, J.C. (2022c), "Pluralia Tantum et verbes supports", *NeoPhilologica*, 34, 1-18.
- Anscombe, J.C. (2023), "Causalité et réactivité dans le fonctionnement du marqueur de discours *je me suis laissé dire*. Application à l'analyse contrastive français/espagnol", *Linguisticae Investigationes*, 46.2, pp. 179-199.
- Anscombe, J.C. (2024), "Le suffixe *-ido* : noms de cris et de bruits naturels en espagnol. Quelques remarques", *Cahiers de lexicologie*, 1, n° 124 : 11-46.
- Badiou-Monferran, C. (2004), " Le statut des expressions figées dans les dictionnaires monolingues de langue française au XVII^e siècle", *Littérature classique*, 50 : 139-165.
- Balibar-Mrabti, A. (2004), "Lexique-Grammaire et extensions lexicales. Note sur le semi-figement ", *Lingvisticae Investigationes Supplementa*, 24, Benjamins, Amsterdam Philadelphia : 23-29.
- Balibar-Mrabti, A. (2005), "Semi-figement et limites de la phrase figée », *Linx*, n° 53, 33-54.
- Balibar-Mrabti, A. (2011), "Ellipse, figement, traduction. *Tomber de Charybde en Scylla*", *Le figement linguistique : la parole entravée*, J.C. Anscombe et S. Mejri (dirs.), Honoré Champion Editeur, Paris : 267-279.
- Balibar-Mrabti, A. (2020), "Forger des phrases simples dans un dictionnaire de langue générale. *Les Verbes français* : un exemple à partir des emplois figurés", *Linx*, n° 20, 1-23.
- Bally, C. (1921), *Traité de stylistique française*, Klincksieck, Paris.
- Barta, P. (2005), "Au royaume des proverbes, les détournements sont rois", *Paremia*, n° 14, 139-152.
- Ben Amor Ben Hamida (2015), "La phraséologie : entre variante(s) et défigement", in *Fraseología, Didáctica y Traducción*, Pedro Mogorron Huerta & Fernando Navarro Domínguez (éds.), Studien zur romanischen Sprachwissenschaft und interkulturellen Kommunikation, Band 101, Peter Lang, Bern : 37-52.
- Benveniste, E. (1966), *Problèmes de linguistique générale*, t. 2, Gallimard, Paris.
- Blanche-Benveniste, C.& Temple, L. (1989), "Ecrire le français parlé", *Le français dans le monde. Recherches et applications*, février-mars, 26-33.

- Blanche-Benveniste, C. & Willems, D. (2016), "Les verbes faibles", in *Encyclopédie Grammaticale du Français*, en ligne : <http://encyclogram.fr>.
- Blanche-Benveniste, C. (2010), *Approches de la langue parlée en français*, Coll. *L'essentiel français*, Ed. Ophrys, Paris.
- Boons, J.P. (1971), "Métaphore et baisse de la redondance", *Langue française*, n° 11, *Syntaxe transformationnelle du français* : 15-16.
- Boons, J.P., Guillet, A., Leclère, C. (1976), *La structure des phrases simples en français. Constructions intransitives*, Droz, Genève-Paris.
- Bréal, M. (1897), *Essai de sémantique : science des significations*, Librairie Hachette & Cie, Paris.
- Buitrago, A. (2002), *Diccionario de dichos y frases hechas*, Espasa Calpe, Madrid, España.
- Chomsky, N. (1965), *Aspect of the Theory of Syntax*,
- Chomsky, N. (1970), *Current Issues in Linguistic Theory*, Mouton, La Haye-Paris.
- Chomsky, N. (1989), "Some notes on economy of derivation and representation", *MIT Working Papers in Linguistics*, 10, 43-74.
- Chomsky, N. (1997), *The Minimalist Program*, N. Chomsky (dir.), The MIT Press, Cambridge, Massachusetts, USA.
- Conenna, M. (2000), "Structures syntaxiques des proverbes français et italiens", *Langages*, 139, 27-38.
- Corbett, G. (2019), "Pluralia tantum nouns and the theory of features: a typology of nouns with non-canonical number properties", *Morphology*, 29, 51-108.
- Craddock, J. & Georges, E. S. (1963-64), "The Hispanic Sound-suffix *-ido*", *RPh* 17, 87-107.
- Croft, W. (2001), *Radical Construction Grammar. Syntactic Theory in Typological Perspective*, Oxford University Press, New-York.
- Danlos, L. (1981), "La morpho-syntaxe des expressions figées", *Langages*, 15, n° 63 : 53-74.
- Danlos, L. (1988), "Lexique-grammaire des expressions figées. Présentation", *Langages*, 23, n° 90 : 5-6.
- Delacourt, F. (1996), *Proverbes, dictons, et citations*, Editions de Vecchi, Paris.
- Dictionnaire de Trévoux* (1750), *Dictionnaire universel françois et latin contenant la signification et la définition tant des mots de l'une et l'autre langue*, 4^{ème} édition, 6 vol., Paris.
- Dournon, J. Y. (1986), *Le dictionnaire des proverbes et dictons de France*, Le Livre de Poche, Hachette, Paris.

- Dubois-Charlier, F. (1972), "La sémantique générative : une nouvelle théorie linguistique ?", *Langages*, 7, n°27 : 5-77.
- Fillmore, C.J. (1968), "The case for case", in *Universals in Linguistic Theory*, Bach & Harms (eds.), New-York: 1-88.
- Fillmore, Ch., Kay P. & Kay O'Connor, M. (1988), "Regularity and idiomacticity in grammatical constructions": the case of *let alone*", *Language*, 64: 501-38.
- Fournier, J.M. (2011), "Notes sur la question du figement dans les descriptions du français à l'âge classique", in *Le figement linguistique : la parole entravée*, Jean-Claude Anscombe & Salah Mejri (éds.), Honoré Champion Editeur, Paris : 299-311.
- Fradin, B. (1984), "Anaphorisation et stéréotypes nominaux", *Lingua*, 64, 325-369.
- François, J. (2008), *Les grammaires de construction : un bâtiment ouvert aux quatre vents*, Crisco, juillet 2008, 1-19.
- Fraser, B. (1970), "Idioms within a Transformational Grammar", *Foundations of Language*, 6: 22-42.
- Freeman, W. (1986), *A Concise Dictionary of English Idioms*, Hodder & Stoughton, London-Sidney-Auckland-Toronto.
- Galmiche, M. (1972), "Représentation sémantique et insertion lexicale", *Langages*, 7, n°27 : 78-126.
- García Page, M. (1999) "Expresiones fijas idiomáticas, semiidiomáticas y libres", *Cahiers du Prohemio*, n° 3: 95-109.
- García Page, M. (2008), *Introducción a la fraseología española. Estudio de las locuciones*, Colección Autores, Textos y Temas, Anthropos, Barcelona.
- García Page, M. (2011), "Aspects sémantiques de la comparative proverbiale du type fuerte como un toro", in *Le figement linguistique : la parole entravée*, Jean-Claude Anscombe & Salah Mejri (éds.), Honoré Champion Editeur, Paris : 127-141.
- García Page, M. (2015), "Relaciones entre la morfología y la fraseología", in *Fraseología y Paremiología: enfoques y aplicaciones*, Venda Durante (ed.), *Biblioteca fraseológica y paremiológica*, «Serie Monografías», n 5, 11-22.
- Gilmont, J.F. (2004), *Une introduction à l'histoire du livre et de la lecture : Du manuscrit à l'ère électronique*, Liège, Céfal.
- Giry-Schneider, J. (1978), *Les nominalisations en français. L'opérateur « faire » dans le lexique*, Droz, Genève-Paris.
- Goddard, C., Peeters, B. (2006), "The Natural Semantic Metalanguage (NSM) approach: An overview with reference to the most important Romance languages", in *Semantic Primes and Universal Grammar. Empirical evidence from the Romance languages*, Bert Peeters (ed.), Benjamins, Amsterdam-Philadelphia: 13-39.
- Goldberg, A. E. (1995), *Constructions. A Construction Grammar Approach to Argument Structure*, The University of Chicago Press, Chicago and London.

- Goldberg, A. E. (2006), *Constructions at Work. The Nature of Generalization in Language*, Oxford University Press, Oxford-New-York.
- Goldberg, A.E. & Jackendoff, R. (2004), "The English Resultative as a family of constructions", *Language*, 80: 532-568.
- Gómez Jordana, S. (2012), "Les moules proverbiaux en français contemporain", in *La parole exemplaire. Introduction à une étude linguistique des proverbes*. Jean-Claude Anscombe, Bernard Darbord & Alexandra Oddo (éds.), Ed. Armand Colin, coll. *Recherche*, Paris : 114-132.
- Greciano, G. (1983), *Signification et dénotation en allemand. La sémantique des expressions idiomatiques*, Klincksieck, Paris.
- Grossman, F. & Tutin, A. (2003), *Les collocations : analyse et traitement*, Ed. de Werelt, Amsterdam.
- Gross, G. (1988), "Degré de figement des noms composés ", *Langages*, 23, n° 90 : 52-72.
- Gross, G. (1996), *Les expressions figées en français*, Ophrys, Coll. *L'essentiel français*, Paris.
- Gross, G. (2012), *Manuel d'analyse linguistique*, Coll. *Sens et structure*, n° 1370, Presses Universitaires du Septentrion, Villeneuve d'Ascq.
- Gross, M. (1975), *Méthodes en syntaxe*, Hermann, Paris.
- Gross, M. (1976), "Présentation", in *La structure des phrases simples en français. Constructions intransitives*, Jean-Paul Boons, Alain Guillet & Christian Leclère (éds.), C. Droz, Genève-Paris : 7-28.
- Gross, M. (1981), "Les bases empiriques de la notion de prédicat sémantique", *Langages*, 15, n° 63, 7-52.
- Gross, M. (1982), "Sur une classification des « phrases figées » du français", *Revue québécoise de linguistique*, 11, n° 2 : 151-185.
- Gross, M. (1984), "Une famille d'adverbes figés : les constructions comparatives en *comme*", *Revue québécoise de linguistique*, 13, n° 2 : 237-269.
- Gross, M. (1985), "Sur les déterminants dans les expressions figées", *Langages*, 20, n° 79 : 89-117.
- Gross, M. (1988), "Sur les phrases figées complexes du français", *Langue française*, n°77 : 47-70.
- Gross, M. (1993), "Les phrases figées en français", *L'information grammaticale*, n°59 : 36-41.
- Gülich, E. & Krafft, U. (1997), "Le rôle du « préfabriqué » dans les processus de production discursive ", in *La locution entre langue et usages*, M. Martins-Baltar (éd.), ENS Editions, Fontenay Saint-Cloud : 241-276.

- Halloran, John A. (2006), *Sumerian Lexicon. A Dictionary Guide to the Ancient Sumerian Language*, Logogram Publishing: Los Angeles.
- Harris, Zellig H. (1964), "Elementary Transformations", *TDAP* N° 54, University of Pennsylvania: Philadelphia. Reprinted in *Papers in Structural and Transformational Linguistics*, 1970, Reidel: Dordrecht.
- Harris, Zellig H. (1968a), "Du morphème à l'expression", *Langages*, 3, n° 9 : 23-50.
- Harris, Zellig H. (1968b), *Mathematical Structures of Language*, Wiley, New-York.
- Harris, Zellig H. (1981), *Papers on Syntax*, Henry Hiż (ed.), Synthese Language Library, 14, D. Reidel Publishing Company: Dordrecht, Boston, London.
- Hoey M. (2005), *Lexical Priming*, Routledge, London.
- Hornstein, N., Nunes, J. & K. Grohmann (2005), *Understanding Minimalism*, Cambridge University Press, Cambridge.
- Hunston, S. & Francis, G. (2000), *Pattern Grammar. A corpus driven approach to the lexical grammar of English*, *Studies in Corpus Linguistics*, John Benjamins Company, Amsterdam-Philadelphia.
- Ibrahim, A. H. (2003), "Le cadre du lexique-grammaire", *Linx*, 48 : 101-122.
- Jespersen, O. (1924), *The Philosophy of Grammar*, George Allen & Unwin Ltd, London.
- Jakendoff, R. (1990), *Semantic Structures* MIT Press., Cambridge, Mass.
- Katz, J.J., Postal, P.M. (1963), "Semantic interpretation of idioms and sentences containing them", *Quarterly Progress Report*, 70: 275-82.
- Katz, J.J., Postal, P.M. (1964), *An Integrated Theory of Linguistic Description*, MIT Press, Cambridge, Mass.
- Kay, P. & Fillmore, Ch. (1999) "Grammatical constructions and linguistic generalizations: the What's X doing Y? construction". *Language*, 75 : 1-33.
- Kleiber, G. (1999), "Les proverbes : des dénominations d'un type « très très » spécial", *Langue française*, 123 : 52-69.
- Kleiber, G. (2000), "Sur le sens des proverbes", *Langages*, 139, 39-58.
- Kuhn, T.S. (1962), *The Structure of Scientific Revolution*, The University of Chicago Press, Chicago and London.
- Lafleur, B. (1991), *Dictionnaire des locutions idiomatiques françaises*, Duculot, Ottawa.
- Lakoff, G. (1968), "Instrumental Adverbs and the Concept of Deep Structure", *Foundations of Language*, vol. 4, n° 1: 4-29.
- Lakoff, G. (1987), *Women, fire, and dangerous things. What categories reveal about the mind*, Chicago University Press, Chicago.
- Lamiroy, B. (2003), "Les notions linguistiques de figement et de contrainte", *Linguisticae Investigationes*, 26/ 1, 1-14.

- Lamiroy, B. & Klein, J. (2005), "Le problème central du figement est le semi-figement", *Linx*, n° 53, 135-154.
- Lamiroy, B. (2008), "Les expressions figées : à la recherche d'une définition", in *Les séquences figées : entre langue et discours*, P. Blumenthal & S. Mejri (éds.), *Zeitschrift für französische Sprache und Literatur*, Cahier n° 36, F. Steiner, Stuttgart : 85-98.
- Lamiroy, B. (2021), "Les expressions verbales figées", in *La Grande Grammaire du Français*, Anne Abeillé & Danièle Godard (dirs.), Editions *Actes Sud*, Arles : 247-63.
- Langacker, R.W. (1987), *Foundations of Cognitive Grammar. Theoretical Prerequisites*, vol. I, Stanford University Press, Stanford, California.
- Lass, R. (1990), "How to do things with junk: Exaptation in language evolution", *Journal of Linguistics* 26, 79–102
- Legallois, D. & François, J. (2006), *Autour des grammaires de constructions et de patterns*, Crisco, janvier 2006, n° 21.
- Legallois, D. & Patard, A. (2017), "Les constructions comme unités de la langue : illustrations, évaluation, critiques", *Langue française*, n° 194, 5-14.
- Legallois (D.), 2018, "La notion de construction", in *Encyclopédie grammaticale du français (EGF)*, en ligne : encyclogram.fr.
- Lépinette, B. (1991), "Etude du *Tesoro de las dos lenguas* (Paris, 1607) de César Oudin", *Iberoromania* : 27-57.
- Lépinette, B. (1999), "Essai de lexicologie contrastive : les expressions figées *Adj. comme N* en français et en espagnol", in *Traduire. Reprises et répétitions*, A.H. Ibrahim & H. Filali (éds.), Presses Universitaires Franc-Comtoises, Col. *Annales littéraires*, n° 675 : 85-99.
- Lépinette, B. (2011), "La difficile émergence de la notion de *figement* dans l'histoire de la grammaire (XVI^e-XVIII^e siècles)", in *Le figement linguistique : la parole entravée*, Jean-Claude Anscombe & Salah Mejri (éds.), Honoré Champion Editeur, Paris : 317-338.
- Leroy, S. (2007), "Les comparaisons *comme SN* exprimant le plus haut degré », *Travaux linguistiques*, 1, n° 54, 69-82.
- Lis, M. & Barbier, M. (1980), *Dictionnaire du gai parler*, Mengès, Paris.
- Litré, E. (1872), *Dictionnaire de la langue française*, Editions du Cap, Monte-Carlo, 1966.
- Luzondo Oyón, A. (2013), "Revisiting Goldberg's Semantic Constraints on the 'way' construction", *RESLA*, 26: 349-364.
- Maloux, M. (1995), *Dictionnaire des proverbes, sentences et maximes*, Coll. « Références », Larousse, Paris.
- Mathieu, Y.Y. (2003), "La Grammaire de Construction", *Linx*, 48, 43-56.

- McCawley, J.D. (1968a), "The role of Semantics in a Grammar", in *Universals in Linguistic Theory*, Bach & Harms (eds.), New-York: 125-170.
- McCawley, J.D. (1968 b), "Lexical Insertion in a Transformational Grammar without Deep Structure", in *Proceedings from the Annual Meeting of the Chicago Linguistic Society, Chicago Linguistic Society*, vol. 4, n° 1 : 71-80.
- Mejri, S. (1997a), *Le figement lexical. Description linguistique et structuration sémantique*, Publications de la Faculté des Lettres de la Manoubia, Série *Linguistique*, n° X.
- Mejri, S. (1997b), "Défigement et jeux de mots", *Etudes linguistiques*, 3, Tunis : 75-92.
- Mejri, S. (2004), "Polysémie et polylexicalité", *Syntaxe et sémantique*, 1, n° 5, *Polysémie et polylexicalité*, Presses universitaires de Caen : 18-30
- Mejri, S. (2005), "Figement absolu ou relatif : la notion de degré de figement", *Linx*, 53 : 183-96.
- Mejri, S. (2011), "Figement, collocation et combinatoire libre", *Le figement linguistique : la parole entravée*, Jean-Claude Anscombe & Salah Mejri (dirs.), Honoré Champion Editeur, Paris : 63-77.
- Mel'čuk, I. (1995), "Phrasemes in Language and Phraseology in Linguistics", in *Idioms: Structural and Psychologiccal perspectives*, M. Everaert & E.-J. Erlbaum (ed.), Hillsdale.
- Mel'čuk, I. (1997), "Vers une linguistique Sens-Texte", *Leçon Inaugurale*, Collège de France, en ligne.
- Mel'čuk et alii (1884-1999), *Dictionnaire Explicatif et Combinatoire du français contemporain. Recherches lexico-sémantiques*, I-IV, Presses de l'Université de Montréal, Montréal.
- Mel'čuk, I. (2006), *Semantics. From Meaning to Text*, vol. I, Studies in Language Companion Series, 129, John Benjamins Publishing Company, Amsterdam/Philadelphia.
- Mel'čuk, I. (2011), "Phrasèmes dans le dictionnaire", in *Le figement linguistique : la parole entravée*, Jean-Claude Anscombe & Salah Mejri (éds.), Honoré Champion Editeur, Paris : 41-61.
- Mel'čuk, I. (2015), "Un modèle linguistique fonctionnel : le modèle Sens-Texte", Conférence, INALCO (22/10/2015), en ligne.
- Mogorrón Huerta, P. (2008), *Analyse du figement et de ses possibles variations dans les constructions verbales espagnoles*, *Linguisticae Investigationes*, Amsterdam/ Philadelphia : John Benjamins.
- Montreynaud, F., Pierron, A. & Suzzoni, F. (1989), *Dictionnaire de proverbes et dictons*, Le Robert, « Les usuels du Robert », Paris.
- Montoro del Arco, E.T. (2006), *Teoría fraseológica de las locuciones particulares. Las locuciones prepositivas, conjuntivas y marcadoras en español*, Peter Lang, Frankfurt am Main.

- Navarro Domínguez, F. (2000), *Analyse du discours et des proverbes chez Balzac*, L'Harmattan, Paris.
- Norrick, N.R. (1985), *How proverbs mean: semantic studies in English proverbs*, Mouton, Berlin.
- Nunberg, G. (1878), *The pragmatics of reference*, Indiana University Linguistic Club, Bloomington.
- Nunberg, G., Sag, I.A., Wasow, T. (1994), "Idioms", *Language* 70, 3: 491-538.
- Oudin, C. (1607), *Tesoro de las dos lenguas española y francesa, Thresor des deux langues françoise et espagnole*, Paris, Marc Orry.
- Oudin, C. (1640), *Curiositez françaises, pour supplément aux dictionnaires*, Paris, Chez Antoine de Sommaville.
- Palma, S. (2007), *Les éléments figés de la langue*, Ed. L'Harmattan, Coll. *Langue et parole*, Paris.
- Pamies, A. (2005), "La comparación estereotipada en español y en francés", en *La creatividad en el lenguaje: colocaciones idiomáticas y fraseología*, JdD. Luque y A. Pamies (eds.). Método, Granada : 469-484.
- Panckoucke, J. (1750), *Dictionnaire des proverbes françois, et des façons de parler comiques, burlesques et familières, etc. avec l'explication, et les étymologies les plus avérées*, à Francfort et à Mayence, chez François Varrentrapf.
- Pérez-Rioja, José A. (1997), *Modismos del español*, Gráfica Cervantes, S.A.: Salamanca, España.
- Pierron, A. (1997), 1997, *Dictionnaire des dictions*, 1997, Marabout, Alleur (Belgique).
- Polguère, A., Mel' čuk, I. (2006), "Dérivations sémantiques et collocations dans le DICO/LAF, *Langue française*, 2, n° 150 : 66-83.
- Polguère, A. (2007), "Soleil insoutenable et chaleur de plomb : le statut lingusitique des greffes collocationnelles", in M.-C. L'Homme et S. Vandaele (dirs), *Les approches lexicographiques et terminologiques sont-elles compatibles?*, Les Presses de l'Université d'Ottawa, Ottawa, 247-291.
- Polguère, A. (2016), *Morphologie et sémantique lexicale. Notions fondamentales*, Les Presses de l'Université de Montréal, Coll. *Paramètres*.
- Quiroga, P. (2005), *Fraseología italo-española*, Granada Lingüística, Granada.
- Rey, A. (2010), *Dictionnaire Historique de la langue française*, Alain Rey (dir.), *Les Dictionnaires Le Robert*, Paris.
- Rey, A. & Chantreau, S. (2003), *Dictionnaire des expressions et locutions*, Coll. *Les usuels*, Le Robert, Paris.
- Ross, J.R. (1967), *Constraints on variables in Syntax*, M.I.T. dissertation, M.I.T, Dept. of Modern Languages and Linguistics.
- Ruiz Gurillo, L. (2001), *Las locuciones en el español actual*, Arco/Libros S.L., Madrid.

- Ruwet, N. (1967), *Introduction à la grammaire générative*, Librairie Plon, Paris.
- Ruwet, N. (1983), "Du bon usage des expressions idiomatiques dans l'argumentation en syntaxe", *Revue québécoise de linguistique*, 13 : 9-145
- Ruwet, N. (1991), *Syntax and Human Experience*, Studies in Contemporary Linguistics, The University of Chicago Press.
- Sag, I.A. & Wasow, T. (1999), *Syntactic Theory: A Formal Introduction*, CSLI Publications, Stanford.
- Sainéan, L. (1920), *Le langage français au xixe siècle : facteurs sociaux, contingents linguistiques, faits sémantiques, influences*, E. de Boccard, Paris.
- Schapira, Ch. (1999), *Les stéréotypes en français : proverbes et autres formules*, Paris, Ophrys.
- Simatos, I. (1997), "Expression idiomatique, expression figée, prédicat complexe", in B. Fradin et J. M. Marandin, éds., *Mots et grammaires*, Didier Eruditio, Paris : 89-127.
- Sinclair, J.M. (1991), *Corpus, Concordance, Collocation*, Oxford University Press,
- Sinclair, J.M. (2004), *Trust the text*, Routledge, London.
- Sommant, M. (2003), *Dictionnaire des locutions et expressions*, Ed. Pocket, Paris.
- Svensson, M.H. (2004), *Critères de figement. L'identification des expressions figées en français contemporain*, Dept. Of Modern Languages/French, Umeå Monograph 15, Umeå University : 197 pp.
- Tamba, I. (2011), "Sens figé : idiomes et proverbes", in *Le figement linguistique : la parole entravée*, Jean-Claude Anscombe & Salah Mejri (éds.), Honoré Champion Editeur, Paris : 110-126.
- Tamba, I. (2014a), "Sémantique lexicale et phrastique : le sens compositionnel revisité", *Cahiers de lexicologie*, 2, n° 105 : 99-116.
- Tamba, I. (2014b), "Du sens littéral au sens compositionnel des proverbes métaphoriques : un petit pas métalinguistique", *Res per Nomen IV, Les théories du sens et de la référence*, E. Hilgert, Silvia Palma, Pierre Frath & René Daval (eds.), Epure, Reims : 501-516.
- Vaguer, C. (2005), "Pédaler dans la semoule. Approches des constructions verbales figées de structures 'V dans GN'", *Linx*, 53 : 231-245.
- Vitri, S. (1990), "On some comparative frozen sentences in Italian", *Lingvisticae Investigationes*, 14, n° 1: 149-174.
- Wierzbicka, A. (1996), *Semantics. Primes and Universal*, Oxford University Press. Oxford-New York : 500 pp.
- William, G. (2003), "Les collocations et l'école contextualiste britannique" in F. Grossmann et A. Tutin (éds.), *Les collocations : analyse et traitement*, coll. *Travaux de recherches en linguistique appliquée*, Ed. De Werelt , Paris: 33-44.

Wood McGee, M. (1986), *A Definition of Idiom*, Indiana University Linguistics Club, Bloomington.

Zuluaga, A. (1980), *Introducción al estudio de las expresiones fijas*, Peter Lang, Frankfurt am Main.

Zuluaga, A. (1999), "Traductología y fraseología", *Paremia*, n° 8, 537-49.

Revues

Langages (1988), 90, *Les expressions figées*, Laurence Danlos (dir.).

Cahiers du français contemporain (1995), 2, *La locution, entre langue et usages*, Blanche-Noëlle Grunig (dir.).

Linx (2005), n° 53, *Le semi-figement* A. Balibar-Mrabti & C. Vaguer (dirs).

Cahiers de praxématique (2006), 47, *Changements linguistiques : figement, lexicalisation, grammaticalisation*, Michelle Lecolle & Sarah Leroy (dirs.).

Langages (2013), 189, *Vers une extension du domaine de la phraséologie*, Dominique Legallois & Agnès Tutin (dirs.).

Pratiques (2013), 159-160, *Le figement en débat*, Laurent Perrin (dir.).

Études de linguistique appliquée (2017), 186/2, *Figement en mouvement et complexité en devenir : regards sur quelques problèmes liés aux structures défigées et complexes*, M.-H. Viguier & A. Grezka (éds.).

Cahiers de lexicologie (2019), 1, n° 114, *Les phrases préfabriquées : sens, fonctions, usages*, Gaétane Dostie & Agnès Tutin (dirs.).

Cahiers de lexicologie (2020), n° 116, *Variation(s) et phraséologie*, Alexandra Oddo, Bernard Darbord & Jean-Claude Anscombe (dirs.).

Langages (2022), 225, *Aspects de la recherche actuelle en phraséologie*, M. Kauffer & Y. Keromnes (eds.).

Verbum (2023), 45/1, *Phraséologie, image et représentation du sens*, Maurice Kauffer & Yvon Keromnes (dirs.).

Verbum (2025), 47/1, *Phraséologie, diasystème, variation linguistique*, Gaétane Dostie (dir.).

LECTURES CONSEILLÉES

Gross, G. (1996), *Les expressions figées en français*, Ophrys, Coll. *L'essentiel français*, Paris.

Legallois, D. & Patard, A. (2017), "Les constructions comme unités de la langue : illustrations, évaluation, critiques", *Langue française*, n° 194, 5-14.

Mel'čuk, I. (1997), "Vers une linguistique Sens-Texte", Leçon Inaugurale, Collège de France, en ligne.

Polguère, A. (2016), *Morphologie et sémantique lexicale. Notions fondamentales*, Les Presses de l'Université de Montréal, Coll. *Paramètres*.

Simatos, I. (1997), "Expression idiomatique, expression figée, prédicat complexe", in B. Fradin et J. M. Marandin, éds., *Mots et grammaires*, Didier Erudition, Paris : 89-127.

Le figement linguistique : la parole entravée (2011), Jean-Claude Anscombe & Salah Mejri (éds.), Honoré Champion Editeur, Paris.