

LES EXPRESSIONS IDIOMATIQUES: DESCRIPTION THÉORIQUE ET TRAITEMENT DICTIONNAIRIQUE (L'EXEMPLE DES DICTIONNAIRES MONOLINGUES FRANÇAIS)¹

Le présent article se propose d'examiner le traitement des expressions idiomatiques dans un corpus de plusieurs dictionnaires monolingues du français en un volume et de suggérer quelques améliorations de ce traitement. L'accent est mis tout particulièrement sur les besoins spécifiques d'un utilisateur étranger. Avant d'entamer une analyse lexicographique proprement dite, il nous semble indispensable de faire le point sur le statut linguistique des expressions idiomatiques et d'en proposer une typologie.

1. LA PLACE DE LA PHRASÉOLOGIE AU SEIN DE LA LINGUISTIQUE

1.1. Depuis une vingtaine d'années, la phraséologie a le vent en poupe². Le nombre de monographies et de colloques, sans compter les articles, consacrés aux recherches théoriques et/ou appliquées des expressions idiomatiques n'a cessé d'augmenter. La grammaire générative et transformationnelle, qui dominait la linguistique des années soixante, soixante-dix et quatre-vingt³, n'a guère porté d'intérêt aux unités phraséologiques (*locutions, idiotismes, expressions figées, clichés, collocations*). Au nom de la «créativité langagière», les linguistes soutenaient que le locuteur jouissait d'une liberté de choix lexicaux presque absolue, les seules contraintes étant celles de la grammaticalité. À la différence de ce point de vue sur le fonctionnement du langage, certains linguistes contemporains qui se sont basés d'avantage sur l'observation de l'usage réel (s'appuyant principalement sur l'exploitation des corpus électroniques ou travaillant dans l'enseignement de langues secondes ou étrangères) essaient de démontrer que les contraintes de la grammaticalité en soi ne suffisent pas à expliquer de nombreux phénomènes linguistiques et discursifs qui relèvent du *figement*. Par exemple, pourquoi les syntagmes *choix heureux* et *conseil précieux* sont-ils acceptables pour un locuteur natif du français, mais non pas **choix précieux* et **conseil heureux*, d'ailleurs parfaitement grammaticaux⁴? De même, pourquoi le choix des composants dans *pren-*

¹ Cet article résume le mémoire de 3^e cycle que l'auteur a soutenu à la Faculté des Lettres de l'Université de Ljubljana en avril 2001 devant le jury: prof. dr. Ada Vidovič Muha (présidente), prof. dr. Dušan Gabrovšek (membre) et prof. dr. Vladimir Pogačnik (directeur de recherche).

² Citons à l'appui de notre remarque introductive A. P. Cowie: «*Following a steady growth of scholarly interest and activity over the past twenty years, chiefly in Western Europe, but also in the USA, phraseology has now become the major field of pure and applied research for Western linguists that it had, much earlier, for scholars in the former Soviet Union and other countries of Eastern Europe.*» (Cowie éd., 1998, p. 1)

³ Au moins en Europe de l'Ouest et aux États-Unis. Cf. Béjoint, 2000, pp. 208-211.

⁴ Les traits de sélection – qui dans la grammaire générative permettent de décrire la grammaticalité des relations syntagmatiques – des deux syntagmes sont identiques: [- concret], [+ comptable].

dre des vessies pour des lanternes, est-il limité aux seules unités lexicales qui y figurent? Ne pourrait-on pas remplacer les deux substantifs (*vessie* et *lanterne*) par deux autres (p. ex. *mouche* et *éléphant*) susceptibles d'exprimer le sens identique («commettre une grossière méprise»)? Pourquoi les lexèmes *prétentaine*, *prou* ou *go* ne s'emploient-ils pas en français contemporain en dehors des expressions *courir la prétentaine*, *peu ou prou* et *tout de go*? Dans ses choix lexicaux, le locuteur n'est pas déterminé uniquement par des relations entre la langue et la réalité extralinguistique et par des contraintes stipulées dans les règles générales de la grammaire d'une langue⁵. Très souvent, le choix d'une unité lexicale entraîne ou impose le choix obligatoire ou, au moins, préférentiel d'une autre unité lexicale ou d'une certaine construction syntaxique. Pour exprimer qu'un choix a été couronné de succès, le locuteur du français dispose d'une liste très limitée d'unités lexicales, en tête de laquelle figure l'adjectif *heureux* comme le choix le plus «idiomatique» – les adjectifs *favorable*, *chanceux*, *bienfaisant*, *avantageux* ou *précieux*, par exemple, étant pour la majorité des locuteurs natifs inacceptables.

1.2. Aujourd'hui, les linguistes ne peuvent plus nier qu'une partie importante de la production discursive (orale ou écrite) est formée de différents types de «séquences préfabriquées» dont la complexité varie considérablement: elles peuvent comprendre quelques mots seulement (*en bloc*, *à l'aide de*), une phrase (*Petite pluie abat grand vent*, *Je vous en prie*, *Mais où sont les neiges d'antan?*) ou bien un texte complet (faire-part de décès, différentes annonces, rituels de félicitation...). Pour décrire adéquatement la langue, il faut donc postuler l'existence de deux principes opposés: la liberté combinatoire et le figement. La dichotomie des deux principes a été introduite en linguistique par Otto Jespersen dans *La philosophie de la grammaire* (1924, la traduction française en 1971). Elle est aujourd'hui abondamment développée et exploitée par un grand nombre de linguistes (théoriques ou appliqués) qui travaillent sur corpus, notamment par des linguistes anglo-saxons⁶.

Si le principe de la liberté combinatoire dominait pratiquement toutes les grammaires, l'ampleur du second principe, du figement⁷, échappait à une grande partie de la linguistique.

2. LE PHÉNOMÈNE DE FIGEMENT ET LES EXPRESSIONS IDIOMATIQUES

2.1. La définition du figement pose des problèmes ardues en linguistique. Les paramètres qui interviennent dans la caractérisation du phénomène sont extrêmement hétérogènes: ils relèvent à la fois de la sémantique, de la syntaxe, de la stylistique, de la pragmatique, de la psycholinguistique. Cette disparité est peut-être une des raisons pour lesquelles le figement est resté si longtemps le laissé-pour-compte de l'intérêt linguistique⁸.

⁵ J. Rey-Debove (1971, pp. 209-213) oppose les *contraintes de contenu* aux *contraintes de signe*.

⁶ L'approche la plus connue est sans doute celle de John Sinclair (1991) qui parle de l'opposition entre «*open-choice principle*» et «*idiom principle*».

⁷ Pour une notion parallèle, celle de *phraséologisation*, qu'utilise Mel'čuk, entre autres, voir plus loin, § 3.3.

⁸ Toutefois, il ne faut pas aller jusqu'à dire que le phénomène était totalement ignoré!

2.1.1. Un premier critère ne soulève pas de problèmes: les linguistes sont unanimes à considérer qu'on ne peut parler du figement que s'il s'agit d'une séquence de plusieurs unités lexicales dont chacune a, au sein d'un système linguistique donné, une existence autonome. Cela exclut d'emblée les mots construits qui relèvent de la dérivation⁹.

2.1.2. Dans la plupart des cas, le figement est décrit comme un processus qui mène de la syntaxe au lexique:

«Lorsqu'un syntagme a la fréquence et la spécificité d'un monème unique, il tend à être traité comme un monème unique, indissociable. On dit qu'il est figé: un bon à rien.» (Mounin, 1995, p. 139)

2.1.3. L'étude du figement se trouve souvent déplacée sur le terrain de la psycholinguistique – le figement est, dans ce cas-là, analysé comme un phénomène à fondement mémoriel: une séquence figée est un syntagme inscrit durablement en mémoire. Les enquêtes démontrent que l'emploi des séquences polylexicales figées se manifeste, à l'oral, «par la rapidité du débit et par l'absence de pauses et d'hésitations» (Gülich & Kraft, 1997, p. 249).

2.1.4. Un des critères de définition qui revient souvent dans la littérature actuelle sur la phraséologie et semble y revêtir une importance particulière relève du niveau sémantique.

2.1.4.1. Ce critère repose sur le *principe de compositionnalité* attribué normalement à G. Frege. Selon ce principe, le sens du tout est fonction des sens des parties – les séquences polylexicales figées, qui vont à l'encontre de ce principe, sont caractérisées par leur non-compositionnalité. Si le sens d'une phrase comme:

Le cuisinier a goûté la moutarde

est déductible des sens des constituants et des fonctions que ces constituants assument, une suite comme:

La moutarde lui monte au nez

ne peut être comprise que dans son intégralité. Même si le locuteur (surtout étranger) connaît les sens de tous les constituants, il n'est pas à même d'interpréter correctement cette séquence («l'impatience, la colère le gagne»). Nombreux sont aussi les cas où les deux lectures, compositionnelle et non-compositionnelle, sont possibles:

Il est de l'autre côté de la barrière

LECTURE COMPOSITIONNELLE:

LE SENS DE L'ÉNONCÉ =

(Il) + (est) + (de) + (l') + (autre) + (côté) + (de) + (la) + (barrière)

LECTURE NON-COMPOSITIONNELLE:

LE SENS DE L'ÉNONCÉ ≠

(Il) + (est) + (de) + (l') + (autre) + (côté) + (de) + (la) + (barrière)

Mais: = (Il est d'opinions politiques différentes)

⁹ Nous laissons de côté le problème des séparateurs (blanc, trait d'union, apostrophe ou soudure) entre les éléments lexicaux formant une séquence figée.

2.1.4.2. Il ne faut pourtant pas perdre de vue que la non-compositionalité des séquences figées n'est pas la cause, mais, au contraire, le résultat de certains processus bien connus en linguistique diachronique. R. Martin (1997), pour ne citer que lui, distingue deux mécanismes principaux:

- l'enrichissement sémantique – le sens de la séquence polylexicale figée n'est pas entièrement prédictible puisqu'il existe des facteurs pragmatiques (culturels ou extralinguistiques) qui «ajoutent» un supplément de sens à la somme des sens des parties. Par exemple, dire d'une femme qu'elle *attend un bébé* ne signifie pas simplement qu'elle «se tient en un lieu où un bébé doit venir», mais qu'elle «est enceinte».
- la démotivation étymologique: la non-compositionalité résulte de la perte de la motivation en synchronie. «Historiquement, la non-compositionalité n'existe pas» (R. Martin, 1997, p. 300).

2.1.5. Un autre critère concerne le comportement syntaxique des expressions figées – il s'agit notamment du blocage des propriétés transformationnelles. Le phénomène a fait l'objet de sérieuses recherches dans la grammaire transformationnelle anglo-saxonne. Pour le français, nous nous référons essentiellement aux travaux accomplis dans le cadre théorique *lexique-grammaire* (G. Gross, 1996).

Une phrase librement construite peut subir un certain nombre de transformations. Prenons comme exemple une construction comportant un verbe transitif direct suivi d'un complément (*Le professeur a vu le film*)¹⁰:

- la passivation: *Le film a été vu par le professeur*
- la pronominalisation: *Le professeur l'a vu*
- la relativisation: *Le film que le professeur a vu*
- le détachement: *Ce film, le professeur l'a vu*
- l'expansion: *Le professeur a vu un film captivant*

Il n'en va pas de même pour les séquences figées (*Le professeur sucre les fraises*):

- **Les fraises sont sucrées par le professeur*
- **Le professeur les a sucrées*
- **Les fraises que le professeur a sucrées*
- **Ces fraises, le professeur les a sucrées*
- **Le professeur sucre des fraises très parfumées.*

2.1.6. Le critère suivant concerne le rapport de la séquence figée à la référence. R. Martin (1997, pp. 300-302) parle de la «saturation intensionnelle», G. Gross (1996, pp. 13-15) de la «non-actualisation des éléments». Comparons, à titre d'illustration, les deux lectures, compositionnelle et non-compositionnelle, de

Le candidat a pris une veste

¹⁰ G. Gross (1996, pp. 12-13) observe que toutes les transformations ne s'appliquent pas de façon systématique à toutes les relations verbe transitif direct-compléments.

La lecture compositionnelle permet l'actualisation du complément *veste*, puisque celui-ci se rapporte à un objet de la réalité extralinguistique référentiellement identifié:

Le candidat a pris une [cette/sa/votre] veste

L'actualisation n'est pas possible dans la séquence figée («le candidat a subi un échec») où le substantif ne réfère à aucun objet réel. En pareil cas¹¹, le substantif ne peut recevoir, en général, qu'une seule forme de détermination: *prendre une veste* (**prendre la/cette veste*), *être à la mode* (**être à une/sa mode*). Le figement peut être également signalé par l'absence d'article: *avoir peur, prendre racine, prendre quelqu'un à témoin, à côté de, à condition que*.

2.1.7. Un critère qui s'ajoute aux précédents touche à l'axe paradigmatique. Un syntagme comme *regarder une montagne* se prête, en position de complément d'objet, à un nombre illimité de substitutions paradigmatiques (synonymiques ou quasi-synonymiques: *regarder un mont, une colline, un pic, un sommet, un massif de montagnes*), tandis que dans des suites figées, les possibilités de substitutions sont bloquées ou au moins considérablement diminuées.

prendre le taureau par les cornes ≈ **prendre le bouc/le bœuf par les cornes*
à bout de force ≈ **à fin de force*
un court-circuit ≈ **un bref-circuit*

2.2. Les critères dégagés se combinent de façons imprédictibles, et rares sont les suites figées qui satisfont à tous les critères énumérés. Telles sont, par exemple, les suites: *c'est la fin des haricots, prendre des vessies pour des lanternes* ou *chambre à air* dont les sens sont opaques et qui ne se prêtent à aucune substitution ou modification¹². D'autres suites, en revanche, permettent certaines variations:

Remporter/ramasser/prendre une veste
Il cherche noise/des noises à ses copains
dans le but (évident) de
Il a faim → *La faim qu'il a*

2.3. À cela s'ajoute la gradualité des critères: le sens d'une suite figée peut être plus ou moins opaque et la suite peut se prêter aux modifications et aux substitutions d'étendue variable. Le figement connaît des degrés et n'est donc pas une «variable booléenne». L'espace qui sépare les suites libres, construites selon les règles générales de la grammaire, des suites complètement figées doit être envisagé comme continuum. À l'un des pôles se situent les constructions dont la combinatoire n'est presque pas restreinte: par exemple, les constructions avec les verbes qui admettent une grande va-

¹¹ R. Martin (1997, p.301) parle de «l'usage intentionnel» du substantif.

¹² Sauf les modifications qui concernent la conjugaison (*c'était la fin des haricots*) et les formes au pluriel (*chambres à air*).

riété de compléments d'objet (*regarder, voir*). Plus on s'éloigne du pôle et plus les contraintes pèsent: le champ de compléments d'objet directs possibles pour le verbe *tuer* (au sens propre) est déjà plus restreint puisqu'il ne comprend que les animés¹³. Il l'est même plus réduit dans le cas du verbe *écarquiller* qui n'admet que l'objet *yeux*. Au pôle opposé, on a affaire à des suites qui répondent à un maximum de critères listés ci-dessus. En premier lieu, il faut mentionner les mots composés¹⁴: *chou-fleur, chambre à air, chemin de fer, à condition de, sans que, à la mode, à poil, faire preuve*. Viennent ensuite les locutions figurées, au sens opaque, dont les éléments ne se prêtent à aucune modification ou substitution: *prendre le taureau par les cornes, les carottes sont cuites*. Plus loin du pôle se situent des suites figées dont les sens sont, même en synchronie, partiellement motivés par les sens des constituants (*être bien dans sa peau, centre commercial, choix heureux*) ou des suites dont les éléments permettent plus ou moins de flexibilité (*ceindre/coiffer/porter la tiare; prendre un pli → le pli a été pris*).

2.4. Bien que le figement soit un phénomène continu, la linguistique se doit, en tant que science, d'y appliquer une grille discrète et d'y distinguer plusieurs catégories de suites figées¹⁵.

2.4.1. Un des premiers à avoir dégagé une classe intermédiaire entre les séquences libres et les séquences (complètement) figées a été Ch. Bally. Dans son *Traité de stylistique française* (1909), il identifie trois types de combinaisons syntagmatiques:

- les associations libres (ou occasionnelles, passagères)
- les séries phraséologiques (ou groupements usuels)
- unités indissolubles.

Examinons maintenant la définition qu'il donne de ces «cas intermédiaires»:

«Il y a série ou groupement usuel lorsque les éléments du groupe conservent leur autonomie, tout en laissant voir une affinité évidente qui les rapproche, de sorte que l'ensemble présente des contours arrêtés et donne l'impression du "déjà vu".» (Bally, 1920, p. 70)

¹³ Le phénomène est normalement analysé comme violation de *restrictions sélectionnelles* (voir aussi la note 4). Certains auteurs (D. A. Cruse, T. Fontenelle) préfèrent aujourd'hui le terme *préférences* au plus «rigoureux» *restrictions*. Les restrictions (ou préférences) sélectionnelles, qui portent sur la compatibilité de grandes catégories sémantiques ([± concret], [± animé]...), sont parfois opposées aux *restrictions* (ou *préférences*) *collocationnelles*, plus «subtiles» et plus imprédictibles. Les deux types de restrictions recourent chez R. Martin (1997, 293-297) la distinction, à l'intérieur de la catégorie générique de *restrictions sélectionnelles*, entre la *limitation de l'étendue combinatoire* et la *sélectivité normative*.

¹⁴ Riegel & Pellat & Rioul (1998, pp. 548-549) dressent un tableau de six structures de base de mots composés: noms composés, adjectifs composés, verbes composés, adverbes composés, prépositions composées, conjonctions composées. G. Gross (1996, 1997) préfère le terme *locution*: il parle ainsi de *locutions verbales, adjectivales, adverbiales, prépositives, conjonctives*, de *déterminants composés* et de *noms composés*. Il critique, d'ailleurs, l'emploi du terme *nom composé* et se demande pourquoi on ne lui préfère *locution nominale*. «Peut-être ce terme classique [noms composés] n'est-il pas très heureux, car il présente du terrain une vue trop compacte et qui ne rend pas compte de l'extrême diversité du paysage.» (G. Gross, 1997, p. 207)

¹⁵ Notre but n'est toutefois pas d'isoler des catégories «étanches», mais plutôt de reconnaître des catégories prototypiques aux frontières plus floues (sur l'approche prototypique, voir Kleiber, 1990).

Les exemples de ce genre de séries que Bally donne ensuite sont aujourd’hui bien connus: *gravement malade, grièvement blessé, désirer ardemment, aimer éperdument*¹⁶. Dans un *groupement usuel*, on reconnaît sans grande difficulté ce que la linguistique contemporaine nomme *collocation*¹⁷. Parmi les propriétés distinctives que recense Bally (1920, pp. 67-79), nous trouvons:

- caractère usuel d’une série qui dépend de l’usage¹⁸;
- la cohésion des termes qui composent une série est seulement «relative» (la cohésion est absolue dans les unités indissociables);
- le sens d’une série dépend en grande partie des sens des termes constituants (par oppositions aux unités indissolubles);
- l’actualisation des termes est admise dans les séries, mais non pas dans les unités (*avoir l’habitude/avoir coutume*);
- les séries n’admettent pas de substitutions synonymiques (*désirer ardemment, aimer éperdument* ≈ **désirer éperdument, *aimer ardemment*).

2.4.2. La notion de *collocation* a été introduite en linguistique dans les années trente par J. R. Firth, un des fondateurs de l’École Contextualiste anglaise. Elle émane d’une conception contextuelle du sens¹⁹ et sert à caractériser certains phénomènes de cooccurrence lexicale qui relèvent de la compétence linguistique des locuteurs natifs, mais que les règles de grammaire sont inaptes à prédire. La collocation n’est ni un mot composé, ni une locution figée ou figurée – elle est constituée d’unités lexicales apparaissant assez souvent ensemble ou au moins à une faible distance les unes des autres. L’occurrence d’une unité lexicale dans un environnement linguistique est pour Firth²⁰ affaire de probabilité: pour toute unité lexicale, il y a des environnements dans lesquels elle a plus de chance de figurer que dans d’autres.

2.4.3. La tendance contextualiste qui se base sur l’interprétation statistique des collocations promue par J. R. Firth est aujourd’hui dominante en TAL(N)²¹ et chez les lexicographes et les terminographes qui utilisent des corpus électroniques (J. Sinclair, T. Fontenelle, U. Heid).

¹⁶ Dans le second volume de son *Traité* (1909, ²1921), Bally propose une série d’exercices consacrés aux groupements usuels (§ 35-§ 47) qui donnent une idée encore plus nette du phénomène en question. Il n’est pas surprenant d’apprendre que l’auteur d’un tel ouvrage travaillait dans l’enseignement du français langue étrangère.

¹⁷ Certains linguistes, surtout anglo-saxons (A. P. Cowie, P. Howart), préfèrent le terme plus précis de «*collocations restreintes*» (ang. *restricted collocations*).

¹⁸ Les collocations sont le plus souvent considérées (cf. Firth, 1957; Hausmann, 1979; R. Martin, 1997) comme un phénomène qui relève de la *norme* linguistique (par opposition au *système*).

¹⁹ «*The complete meaning of the word is always contextual and no study of meaning apart from a complete context can be taken seriously.*» (Firth, 1957, p. 7).

²⁰ La conception firthienne des collocations a été développée dans les années soixante et soixante-dix par M. A. K. Halliday et T. F. Mitchell.

²¹ TAL(N) est une abréviation de *traitement automatique du langage (naturel)*, qui est d’ailleurs un calque («maladroit», selon certains) de l’anglais «*NLP*» (*Natural Language Processing*). Dans le domaine du TALN,

2.4.3.1. Les collocations sont définies en termes de fréquences²²: comme suites de deux ou plusieurs unités lexicales qui sont fréquemment cooccurentes²³ à l'intérieur d'une fenêtre²⁴. Les linguistes travaillant en TALN tendent à développer des méthodes (ou outils) d'identifications automatiques des collocations en corpus. Ces méthodes reposent, en règle générale, sur différentes mesures statistiques²⁵: une des plus rudimentaires consiste en un simple calcul des fréquences; parmi les méthodes plus «pointues», mentionnons:

- la méthode basée sur la moyenne et la variance de la distance entre un mot donné et des mots cooccurents;
- le «T-test» ou «T-score» (mesure du contraste, dont «Z-score» est une des variantes);
- le « χ^2 test» (le «test khi carré»);
- «Mutual Information» (mesure de l'association).

2.4.3.2. Sans entrer dans le détail, ajoutons toutefois que les résultats obtenus à l'aide de ces méthodes doivent être interprétés avec prudence. En premier chef, les comparaisons des différentes méthodes, même appliquées au même corpus, ont montré que les résultats, souvent, ne se recouvraient que partiellement (Cf. Manning & Schütze 1999, pp. 153-183). D'ailleurs, aucune mesure statistique ne fonctionne bien si les fréquences d'occurrences sont basses²⁶. De plus, ces mesures dégagent des combinaisons de mots qui, selon d'autres définitions (sémantiques, syntaxiques ou pragmatiques) et intuitivement, ne peuvent être considérées comme des collocations²⁷. Comme l'a fait judicieusement remarquer Van der Wouden (1997, p. 34), les méthodes statistiques sont très utiles pour la sélection des «candidats» au statut de collocation, mais n'ont pas de moyens pour séparer, par elles-mêmes, les collocations des autres types de combinaisons lexicales²⁸.

2.4.4. Une autre tendance dans l'étude des collocations regroupe des approches qui visent à la formalisation des collocations et à leur description lexicographique.

2.4.4.1. Il convient d'abord d'examiner les études théoriques de F. J. Hausmann (1979,

les collocations posent des problèmes pour la génération de textes, pour l'encodage, pour la désambiguïsation et pour la traduction automatique.

²² Les définitions qu'on trouve parfois peuvent être très «lâches». Citons, par exemple, J. Sinclair: «*Collocation is the occurrence of two or more words within a short space of each other in a text*». Il ajoute pourtant: «*Collocation can be dramatic and interesting because unexpected, or they can be important in the lexical structure of the language because of being frequently repeated*». (Sinclair, 1991, p. 197)

²³ C'est-à-dire que les mots qui forment les collocations se trouvent à proximité les uns des autres beaucoup plus souvent qu'on aurait pu le prévoir si l'on considère le nombre d'occurrences de chaque mot pris séparément.

²⁴ Pour le repérage des collocations en corpus, la fenêtre (ang. *window*) comporte normalement 3 ou 4 mots de chaque côté du mot focal.

²⁵ Dans le bilan de différentes mesures statistiques, nous nous appuyons essentiellement sur Manning & Schütze (1999, pp. 151- 189).

²⁶ Ce qui explique pourquoi il est indispensable d'utiliser, pour le repérage des collocations, des corpus de grande taille comptant des dizaines de millions de mots.

²⁷ Dans les corpus en langue anglaise (et américaine), ces mesures peuvent extraire aussi des suites comme «*I love you*», «*how do you do*» ou «*an apple*». Cf. Van der Wouden, 1997, pp. 31-34.

²⁸ Généralement, le terme *collocation* a une extension plus grande en TALN qu'en lexicologie et en lexicographie (voir Manning & Schütze, 1999, pp.183-187).

1989). Son approche s'appuie dans un premier temps sur des paramètres de nature syntaxique. La collocation, selon Hausmann, est une «*combinaison caractéristique*» (1989, p. 1010) binaire qui prend, en français, une des formes suivantes:

- substantif + adjectif (*célibataire endurci*)
- substantif + verbe (*la colère s'apaise*)
- verbe + substantif (*retirer de l'argent*)
- verbe + adverbe (*il pleut à verse*)
- adjectif + adverbe (*grièvement blessé*)
- substantif (+ préposition) + substantif (*une bouffée de colère*)

Les collocations, faits de langue et non pas faits de parole, se distinguent à la fois des combinaisons libres (par la «combinabilité restreinte» des constituants) et des locutions (par le non-figement et la non-transparence sémantique). Elles sont pourtant imprédictibles – surtout pour un locuteur étranger en situation d'encodage: parmi plusieurs combinaisons sémantiquement et syntaxiquement possibles, la langue choisit, de façon idiosyncrasique, une seule ou un nombre très réduit de ces combinaisons. Hausmann insiste tout particulièrement sur l'asymétrie sémantique des deux parties de la collocation, ce qui n'est nullement sans conséquence pour la description lexicographique des collocations (voir plus loin)²⁹:

«*Dans la collocation, le statut des deux partenaires combinés n'est pas égal. L'un des partenaires, par ex. le célibataire, est autonome sur le plan sémantique. L'autre partenaire (endurci) ajoute une caractéristique qui ne modifie pas l'identité du caractérisé.*» (Hausmann, 1989, 1010)

Le constituant autonome est nommé *base* (dans les exemples donnés ci-dessus, ce sont *célibataire, colère, argent, pleuvoir, blessé, colère*) et le constituant dépendant (dont le sens précis ne peut être déterminé sans référence à l'autre élément de la collocation) *collocatif*.

2.4.4.2. Une approche novatrice des problèmes que soulèvent les descriptions théorique et lexicographique des collocations a été proposée par le linguiste russe I. A. Mel'čuk. L'approche se situe dans le cadre théorique *Sens-Texte* que Mel'čuk, Zholkovky et Apresjan développent depuis les années soixante. Le but visé par la théorie est la construction d'un modèle fonctionnel, appelé *Modèle Sens-Texte*, susceptible de simuler au mieux la correspondance entre le sens qu'un locuteur veut exprimer et le texte (ou les textes) qui le véhicule³⁰. La théorie prévoit un type particulier de dictionnaire, appelé *Dictionnaire explicatif et combinatoire (DEC)*³¹, qui fait partie intégrante

²⁹ Hausmann (1999) propose même une nouvelle discipline, appelée *semiotaxis*, dont le but serait d'examiner dans quelle mesure les mots ont besoin de contexte pour pouvoir être définis. Il distingue entre les mots *sémiotaxiquement indépendants* ou *autonomes* (p. ex. *thermomètre, célibataire, dormir, mourir*) et les mots *sémiotaxiquement dépendants* (p.ex. *dur, endurci, serrer*). Il serait intéressant de comparer ces notions avec la notion de *dépendance (conceptuelle)*, en ang. (*conceptual*) *dependence*, de R. W. Langacker (1987).

³⁰ Bien que le modèle soit bidirectionnel (*Sens ↔ Texte*), il privilégie la synthèse (l'encodage, la production): *Sens ⇒ Texte*. Cf. Mel'čuk, 1993, pp. 46-47; 1997a, pp. 14-15.

³¹ Pour plus de détails, voir: Mel'čuk et al., 1984, 1988, 1992, 1999; Mel'čuk & Clas & Polguère, 1995.

du modèle, notamment de sa composante sémantique. Le *DEC* est un dictionnaire «expérimental» et «idéalisé» qui vise une description formelle, explicite et exhaustive de tous les moyens lexicaux dont dispose une langue pour exprimer un sens donné: en principe, le *DEC* doit fournir toutes les informations dont un locuteur non-natif a besoin pour construire toutes les expressions correctes et acceptables de n'importe quel concept considéré. Contrairement à la majorité des dictionnaires monolingues, le *DEC* est un dictionnaire orienté principalement vers l'encodage.

2.4.4.2.1. Un des concepts-clés du *DEC*, et du modèle en général, est celui de *fonctions lexicales* (FL). La découverte des FL est due à l'observation que les collocations, ces «bizarreries de l'usage» du type *grièvement blessé* (**grièvement malade*, mais *gravement malade* et *gravement blessé*), *exciter l'intérêt* (**soulever l'intérêt*), *soulever l'enthousiasme* (mais, également, *exciter l'enthousiasme*), se rencontrent très souvent «avec un nombre fort réduit de sens spécifiques – très abstraits et généraux» (Mel'čuk, 1997a, p. 45). Si l'on examine le phénomène dans une perspective de l'encodage, on dirait que pour exprimer, par exemple, le sens «intense/intensément» sous forme d'une collocation, le choix d'un des éléments de la collocation dépend du choix de l'autre: c'est pour cela qu'un *amour est fou*, une *peur bleue*, qu'on *souffre atrocement* et qu'on *nie catégoriquement*³². Un tel sens peut être formalisé par une fonction lexicale (qui ressemble à une fonction mathématique) *f* qui associe à toute unité lexicale *x* (pour laquelle ce sens peut être exprimé) toutes les unités *y* possibles, susceptibles de l'exprimer.

$$f(x) = y$$

L'élément *x*, argument de la fonction *f*, est appelé *mot-clé* et l'élément *y* est nommé *valeur*. Les FL sont universelles et peu nombreuses (à peu près 60)³³: outre les FL standard, le modèle connaît les FL non standard, ainsi que les FL complexes et les configurations des FL.

Les exemples de collocations donnés ci-dessus peuvent être représentés sous forme de la FL Magn (du latin *magnus*, exprimant le sens «intense, à un haut degré, très»):

Magn (*amour*) = *fou*

Magn (*souffrir*) = *atrocement*

Magn (*peur*) = *bleue*

Magn (*nier*) = *catégoriquement*

2.4.4.2.2. Disons dès l'entrée en matière que les concepts de FL et de collocation ne sont pas coextensifs. Les FL couvrent, outre les relations syntagmatiques relevant de la cooccurrence lexicale restreinte³⁴, les relations paradigmatiques, comme la synonymie /Syn (*voiture*) = *automobile*/, la conversion /Conv₂₁ (*effrayer*) = *craindre*/, l'au-

³² Pour Ch. Bally, l'expression de l'intensité représente une partie importante (même une «catégorie à part») des séries usuelles (= collocations). Une autre catégorie importante (appelée *séries verbales*) est constituée par des constructions à verbe support (*prendre une décision, remporter une victoire*). Voir Bally, 1920, pp.70-72.

³³ La liste complète de toutes les FL se trouve au début de chaque volume du *DECFC* (*Dictionnaire explicatif et combinatoire du français contemporain* – Mel'čuk et al., 1984, 1988, 1992, 1999) et dans: Mel'čuk & Clas & Polguère, 1995; Mel'čuk, 1996.

³⁴ Ce sont des cooccurents dont la combinatoire ne peut être déterminé ni par leur sémantisme ni par leurs propriétés syntaxiques.

tonymie /Anti (*respect*) = *irrespect*/, d'hypéronymie /Gener (*armoire*) = *meuble*/, la dérivation syntaxique /S₀ (*présenter*) = *présentation*/, la dérivation sémantique /S₁ (*parler*) = *locuteur*; S₂ (*parler*) = *paroles, propos, discours*; S₃ (*parler*) = *allocutaire, destinataire*/.

2.4.4.2.3. Nous avons déjà évoqué une typologie des FL³⁵ (c'est d'ailleurs la typologie qui est le plus souvent citée), à savoir celle qui distingue les FL *paradigmatiques* des FL *syntagmatiques*. Les FL syntagmatiques visent à formaliser le phénomène de la cooccurrence lexicale restreinte. Parmi les FL syntagmatiques les plus importantes, citons, à titre d'illustration:

- Magn (voir ci-dessus)
- Bon (modificateur adjectival ou adverbial exprimant l'approbation du locuteur)
Bon (*conseil*) = *précieux* Bon (*choix*) = *heureux*
- Oper_i, Func_i, Labor_{ij} (le triplet formalise la notion de *verbe support*³⁶ – le choix de la FL est déterminé par le rôle syntaxique que joue le mot-clé; les indices renvoient aux actants syntaxiques profonds du mot-clé):

Oper_i: son mot-clé est son complément d'objet direct

Oper₁ (*erreur*) = *faire l'erreur* de X (= actant 1) dans Y (= actant 2)/

Oper₂ (*applaudissement*) = *recevoir l'applaudissements* de X (=1) destinés à Y (=2)/

Oper₃ (*ordre*) = *recevoir l'ordre* de X (=1) à Y (=2) de faire qqch. (=3)/

Func_i: son mot-clé est son sujet

Func₀ (*réunion*) = *est en cours*

Func₁ (*aide*) = *vient [de N]*

Func₂ (*danger*) = *menace [N]*

Func₃ (*interdiction*) = *frappe [N]*

Labor_{ij}: son mot-clé est son complément d'objet indirect

Labor₁₂ (*location*) = *donner [N en ~]*

Labor₃₂ (*location*) = *prendre [N en ~]*

- Real_i, Fact_i, Labreal_{ij} (ces FL décrivent des verbes sémantiquement pleins exprimant le sens «réaliser l'objectif inhérent à la chose désignée par le mot-clé» – en ce qui concerne la syntaxe³⁷, c'est-à-dire les rôles joués par le mot-clé et ses actants syntaxiques profonds, Real_i est identique à Oper_i, Fact_i à Func_i et Labreal_{ij} à Labor_{ij}):

Real₁ (*peine*) = *imposer, infliger [ART ~]*

Real₂ (*peine*) = *purger [ART ~]*

Fact₀ (*rêve*) = *se réalise*

Fact₂ (*médecin*) = *reçoit, voit [N]*

Labreal₁₂ (*piège*) = *prendre [N dans ART ~]*

- Son («produire le son typique»)

Son (*éléphant*) = *barrit* Son (*canon*) = *gronde*

³⁵ Pour une revue complète des différentes typologies des FL, voir Alonso Ramos & Tutin, 1996.

³⁶ La notion a été introduite par M. Gross (1981, pp. 16 *ssq.*): les verbes supports sont les verbes sémantiquement vidés (*avoir, être, faire, prendre, perdre, donner, recevoir*) qui ne servent qu'à actualiser («verbaliser») les noms à sens prédicatif (*ordre, reproche, conseil...*).

³⁷ Nous pensons à la syntaxe logique.

2.4.4.2.4. À notre connaissance, les FL constituent le seul outil linguistique qui décrive et classe les collocations, d'une façon systématique et rigoureuse, selon des critères essentiellement sémantiques³⁸. En raison, sans doute, d'une extrême complexité de l'approche de Mel'čuk et des difficultés que soulève la consultation des formalisations très poussées de ses descriptions lexicographiques pour des «non-initiés», les FL ne connaissent presque pas d'applications directes en apprentissage des langues étrangères (voir Mel'čuk, 1993a; Polguère, 2000) ni en lexicographie grand public. Dans la suite, nous signalerons toutefois quelques aspects de leur utilisation possible en lexicographie monolingue.

2.5. Avant de proposer notre typologie d'expressions idiomatiques, il nous reste à aborder un dernier groupe de critères: il s'agit de critères pragmatiques. Pour les décrire, I. Fónagy (1982, 1997) a introduit les notions de *liage* et d'*énoncé lié*. Le liage, type particulier du figement, vise à expliquer le fait qu'il existe des énoncés préfabriqués dont l'emploi dépend étroitement des conditions pragmatiques. Ces énoncés *liés* sont «*presque automatiquement déclanchés par telle ou telle situation caractérisée*» (Fónagy, 1997, p. 132). Les situations récurrentes de notre vie quotidienne donnent lieu à un nombre limité d'énoncés: ce nombre est bien inférieur au nombre d'énoncés que la grammaire (la syntaxe et la sémantique) aurait permis – les contraintes qui entrent en jeu sont de nature pragmatique. Par exemple, au début d'une conversation téléphonique, si un locuteur français ne reconnaît pas la voix de la personne qui désire lui parler, il emploiera à coup sûr la formule: «*C'est de la part de qui?*», les autres énoncés (p. ex. *?A qui je parle?*, *?Qui appelle?* ou *?Qui parle?*) «linguistiquement» possibles étant exclus. Il en va de même pour les autres formules qui jalonnent les communications téléphoniques en français (*Allo, je suis bien chez M. X?*, *Pourriez-vous me passer M. X?*, *Ne quittez pas*) et d'ailleurs pour tous les autres types de communications quotidiennes (*Bonne journée, Bon courage, Je t'/vous en prie, Je le dis comme ça, Ça ne fait rien, Passez donc, Ne bougez pas*).

2.5.1. Sans trop nous risquer dans le «maquis» terminologique, mentionnons que F. Coulmas, pour désigner ce type d'expressions, utilise le terme *routines discursives* (ang. *conversational routines*); Cowie (1994, 1998) et de nombreux autres linguistes anglo-saxons parlent de *formules* (ang. *formulae*); H. Burger (voir Gülich & Kraft,

38 Hausmann (1979), déjà, en vue d'une plus grande systématisation des formes de présentation des collocations, prévoit, à l'intérieur du cadre des constructions syntaxiques, les sous-classes où les collocations sont réparties selon «un système sémantique simple». Les sous-classes sont introduites par des termes génériques, qui rappellent les FL de Mel'čuk. Voici un extrait de l'article consacré au nom *doute*:

«1. [S + v] NAÎTRE, EXISTER: naître, surgir, m'envahit, plane, subsiste, persiste; DISPARAÎTRE: s'évanouir, s'envoler. 2. [v + S] AVOIR: avoir, concevoir, éprouver, il me vient des doutes; FAIRE NAÎTRE: inspirer; EXPRIMER: émettre, formuler; FAIRE DISPARAÎTRE: lever, écarter, éclaircir, dissiper, balayer. [...]».

Les rédacteurs du *BBF* (Benson, M. & Benson, E. & Ilson, R., 1997), de même, ajoutent au classement grammatical des collocations des paramètres sémantiques.

1997, p. 245) a forgé le terme «*idiotismes pragmatiques*», qu'il a remplacé plus tard par «*phraséologismes pragmatiques*» (all. *pragmatische Phraseologismen*). Le terme que nous emploierons pour désigner des séquences contraintes relativement aux critères pragmatiques est celui de *pragmatèmes* (ou *phrasèmes pragmatiques*) que nous empruntons à Mel'čuk (1993a, 1998).

2.5.2. Fónagy observe que le phénomène de liage s'accompagne d'un affaiblissement sémantique des énoncés: il y a un clivage d'ordre «quantitatif» entre la signification littérale (donc compositionnelle) d'un pragmatème (= énoncé lié) et sa signification réelle en discours – les pragmatèmes «signifient» moins. En termes jakobsoniens, nous dirions que ce sont les fonctions conative et phatique qui l'emportent sur la fonction référentielle des pragmatèmes. Comparez, à titre d'exemple, les significations des énoncés libres avec celles des pragmatèmes homonymes (cf. Fónagy, 1982, pp. 52-57; 1997, pp. 141-142):

Tiens la corde! ≈ **Tiens**, par exemple...!

Tu veux? J'en ai un autre. ≈ *Ferme-la, tu veux?*

Si vous voulez, venez! ≈ **Donc, si vous voulez**, c'est vraiment une catastrophe.

2.5.3. C'est justement grâce à cet appauvrissement sémantique que les pragmatèmes peuvent remplir une grande diversité des fonctions dans le discours. Fónagy (1997, pp. 143-146) distingue 5 classes de pragmatèmes:

- *fonctions sémantiques*: l'appauvrissement a pour conséquence la synonymie qui permet à un locuteur d'avoir recours à un large éventail d'énoncés qui se différencient seulement d'après leur valeur stylistique, leurs connotations ou leur degré d'intensité (par exemple, différentes expressions du désaccord: «*Tu dis n'importe quoi*», «*Ne sois pas ridicule!*», «*Je le regrette*», «*C'est une opinion*», «*Vous dites ça?*», «*Écoutez*», «*Moi, je suis contre*», «*Oublie-moi*», «*Vous êtes malade?*», «*Voyez-vous ça!*», «*Quelle idée!*», «*Ça va pas la tête!*»);
- *fonctions discursives*: les pragmatèmes aident à organiser le discours («*Eh bien voilà*», «*Venons-en maintenant à...*», «*Disons-le tout de suite*»);
- *facilitation de la communication*: les pragmatèmes sont un grand réservoir qui permet à un locuteur de trouver à peu de frais une expression conforme à une situation de communication donnée, de meubler les silences gênants de différents clichés («*il y toujours des choses*»), de nuancer ses propos («*Je le dis comme ça*») ou d'éviter de répondre («*Je me comprends*»);
- *fonctions sociales*: les pragmatèmes fournissent des lieux communs qui manifestent et renforcent l'appartenance des locuteurs à des groupes sociaux; parfois ces énoncés dégénèrent jusqu'à créer une sorte de «langue de bois»;
- *conservation de la croyance en la magie verbale*: les pragmatèmes représentent les vestiges de la croyance dans le pouvoir bénéfique (p. ex. les formules de politesse) ou maléfique (les jurons) des mots.

2.5.4. Une partie des linguistes (Fónagy, Mel'čuk ou Cowie) considèrent que les pragmatèmes, normalement, ne dépassent pas le niveau phrastique. Très souvent même, l'étendue du pragmatème est réduite à un seul mot (les tournures elliptiques: *Pardon*, *Affirmatif!*, *Merci*) ou à un «fragment» qui fournit le cadre dans lequel s'insèrent des énoncés variés («*L'ennuyeux c'est que...*», «*Tu peux rigoler, mais...*», «*...pendant que vous y êtes...*», «*C'est pas... qui étouffe*»). D'autres linguistes (Gülich & Kraft, 1997), en revanche, insistent sur la nécessité d'inclure aussi l'étude des modèles de textes ou des «textes stéréotypés»: faire-part de naissance, de mariage, de décès, remerciements, discours de bienvenue, petites annonces, différents types de lettres etc.

3. POUR UNE TYPOLOGIE DES EXPRESSIONS IDIOMATIQUES

Après avoir passé en revue les différentes caractéristiques (d'ordre sémantique, syntaxique, psycholinguistique, statistique, pragmatique) des expressions figées, il est clair que nous devons distinguer plusieurs classes au sein de cette catégorie: une dichotomie qui opposerait les séquences libres à une masse indifférenciée de suites figées serait sans valeur théorique et, comme nous le verrons plus en détail dans la suite, également sans valeur pratique pour la lexicographie.

3.1. Nous avons déjà mentionné que Ch. Bally (voir ci-dessus) a introduit, dans son classement des groupes syntagmatiques selon le degré de «fixité», entre les *groupements passagers* et les *unités indissolubles* (c'est-à-dire, *locutions*) une catégorie intermédiaire appelée *séries phraséologiques* (aujourd'hui communément nommés *collocations*). Ces typologies tripartites, comportant les collocations comme une catégorie «intermédiaire» entre les séquences libres et les séquences (complètement) figées, se retrouvent également dans de nombreux travaux phraséologiques contemporains. Sous l'influence des modèles phraséologiques russes (V. V. Vinogradov, N. N. Amosova), les phraséologues anglo-saxons (A. P. Cowie, 1981) subdivisent souvent la catégorie des *locutions* (ang. *idioms*; c'est-à-dire, la catégorie des séquences complètement figées) en *locutions figurées* et *locutions pures*.

3.2. Depuis Burger (voir Gülich & Kraft, 1997), qui a été l'un des premiers à le proposer, un nombre de linguistes de plus en plus important intègrent les pragmatèmes (ou ses équivalents terminologiques) dans leurs typologies des expressions idiomatiques – Cowie (1994), par exemple, isole 3 grandes classes: *collocations*, *locutions* (ang. *idioms*) et *formules* (ang. *formulae*; il s'agit de nos pragmatèmes).

3.3. Une place à part dans ce bref examen de différents classements revient à I. A. Mel'čuk. Sa typologie, remarquable par son exhaustivité et sa rigueur, se base sur la notion de *phraséologisation*. La phraséologisation (voir Mel'čuk, 1997, p. 393) désigne le phénomène où une suite d'unités lexicales possède des propriétés (sémantiques, phonétiques, combinatoires ou pragmatiques) que l'on ne peut pas dériver, selon

les règles générales d'une langue, des propriétés correspondantes des unités constituantes. Une telle suite d'unités, qui s'oppose à une suite construite exclusivement selon les règles générales d'une langue, est appelée *phrasème*. Mel'čuk distingue:

- phrasèmes sémantiques (l'idiosyncrasie concerne les propriétés sémantiques)
- phrasèmes formels (l'idiosyncrasie concerne les propriétés phonétiques)
- phrasèmes combinatoires (l'idiosyncrasie concerne les propriétés combinatoires)
- phrasèmes pragmatiques ou pragmatèmes (l'idiosyncrasie concerne les propriétés pragmatiques).

Les phrasèmes sémantiques³⁹ se divisent, en plus, en:

- *phrasèmes sémantiques complets* (ou *forts*) – le sens du phrasème n'inclut le sens d'aucun de ses constituants (*tenir la chandelle, pont aux ânes*);
- *phrasèmes sémantiquement faibles* ou *semi-phrasèmes* (= collocations) – le sens du semi-phrasème inclut le sens d'un de ses constituants, tandis que l'autre constituant, qui n'est pas choisi librement, le plus souvent, ne garde pas son sens originel (*décocher un compliment, peur bleue*);
- *quasi-phrasèmes sémantiques* – le sens du phrasème inclut les sens de tous ses constituants, puis, en surplus, des traits sémantiques imprévisibles (*appeler à l'aide, centre commercial*).

3.4. Notre typologie oppose aux syntagmes librement construits, improvisés à chaque fois selon les règles standard d'une langue, la catégorie des *expressions idiomatiques*, inscrites en mémoire et qui, par conséquent, ne se désagrègent pas après l'énonciation. Nous y distinguons trois classes de base:

- les *collocations* sont des expressions récurrentes formées de deux ou (rarement) de plusieurs unités lexicales syntaxiquement reliées, dont la cooccurrence se révèle imprédictible à l'encodage: le choix d'un des éléments de la collocation (appelé *collocatif*) est «contraint» par le choix de l'autre (dit *base*), qui est, lui, choisi librement en fonction de son sens (*dormir profondément, choix heureux, poser une question, meute de chiens*)⁴⁰. Au niveau du décodage, nous pouvons noter que le sens de l'unité qui est la base de la collocation reste inchangé, tandis que le sens de l'autre élément, du collocatif, subit d'importantes modifications contextuelles. Par exemple, dans *rompre le silence*, la base *silence* garde son sens originel, ce que nous ne pouvons pas dire pour le collocatif *rompre*, qui prend le sens «faire cesser par un son, une parole» (et cela uniquement dans le contexte de la collocation);
- les *pragmatèmes* sont des préconstruits langagiers dont l'emploi dépend étroitement des critères pragmatiques: ils sont figés par rapport à des situations caractérisées.

³⁹ Un examen détaillé de la typologie de Mel'čuk dépasserait largement le cadre du présent article. Nous gliserons sur les phrasèmes formels et combinatoires (voir Mel'čuk, 1997, pp. 293-404) ainsi que sur les pragmatèmes dont il a été déjà question ci-dessus.

⁴⁰ Les bases des collocations sont en gras.

Souvent, leur sens est transparent, mais appauvri. La catégorie paraît à première vue très hétérogène puisqu'elle embrasse salutations (*A bientôt*), formules de politesse (*Je t'/vous en prie, Veuillez agréer, Monsieur, l'expression de mon respect*), formules accompagnant différentes phases de la conversation (*Que voulez-vous que je vous dise, Dites donc, Tu veux, Eh bien voilà*), «lieux communs» (*La vie, c'est la vie, Finalement, rien n'est joué*), plaisanteries préfabriquées (*Tu a trouvé ton permis de conduire dans une pochette-surprise?*), méta énoncés (*Pour me servir d'un terme cher à..., Comme dirait l'autre*), proverbes (*Tel père, tel fils*), citations (*La réforme, oui; la chienlit, non*), clichés (*J'en passe et des meilleurs, Silence, on tourne*). On pourrait ajouter encore des modèles de textes entiers: lettres, petites annonces, faire-part, remerciements...

- les *phrasèmes* sont des expressions polylexicales qui connaissent un haut degré de figement: ils n'admettent presque pas de modifications, et leur sens est dans une large mesure idiosyncrasique relativement au sens de leurs constituants. La catégorie recouvre les notions linguistiques «traditionnelles» de locutions (*prendre des vessies pour des lanternes, de peur de, à poil, sans que*) et de mots composés (*chambre à air, fait divers, mouvement uniformément accéléré*).

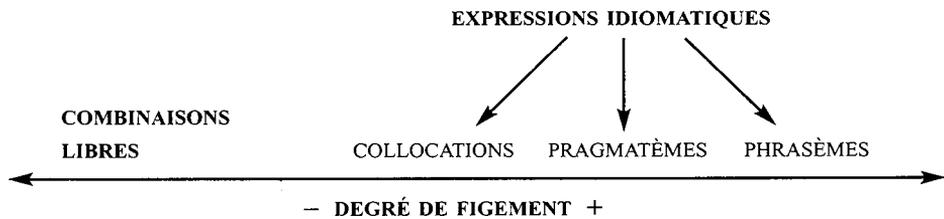
3.5. Comme notre analyse du phénomène de *figement* ainsi que nos définitions de différents types d'expressions idiomatiques ont dû suffisamment le mettre au jour, nous ne considérons pas les classes majeures de notre typologie comme des catégories discrètes, «étanches» et absolues, mais, plutôt, comme des catégories prototypiques (voir n. 15). Les facteurs de figement sont graduels, et leur combinatoire est souvent imprévisible et sujette à la variabilité. Chaque catégorie connaît des exemples centraux (donc prototypiques), qui possèdent le plus grand nombre de propriétés énumérées dans les définitions, et des exemples qui se situent plus à proximité des frontières et qui, souvent, soulèvent d'insolubles problèmes de catégorisation. La ligne de démarcation est particulièrement «floue» entre les collocations et les phrasèmes d'une part, et les collocations et les syntagmes libres de l'autre. Par exemple, faut-il reconnaître en *avoir faim, faire froid* ou *faire peur* des collocations (de même qu'en *faire une remarque, donner un ordre, donner une conférence*) ou bien des phrasèmes, étant donné l'absence de déterminants et le blocage d'un certain nombre de transformations? Les phraséologues (Cowie, 1994; Gabrovšek, 1998), en effet, soulignent souvent que les collocations sont des combinaisons sémantiquement transparentes au décodage (*faire la grève, donner un ordre, joie immense*) – cependant, il ne faut pas perdre de vue que le changement de sens que subit le collocatif dans le contexte d'une collocation peut engendrer un certain degré d'opacité: par exemple, *entamer des négociations, rompre le silence, choix heureux, reproche sanglant*⁴¹. En ce qui concerne l'imprédictibilité

⁴¹ Dans le cas d'un locuteur étranger, la transparence ou l'opacité d'une collocation dépendent également de l'interférence possible de sa langue maternelle: par exemple, pour un Slovène, les collocations françaises

des collocations, ce qui les distingue des combinaisons libres, nous devons noter qu'elle n'est pas absolue non plus – par exemple, le verbe *éprouver* peut figurer comme collocatif avec tous les noms dénotant les émotions: *éprouver de la joie, de la haine, de la colère, de la tendresse, de l'affection, de l'amour, de la honte, un regret...* Ces combinaisons, sont-elles des collocations ou des combinaisons libres⁴²?

Le critère de transparence sémantique (c'est-à-dire de la compositionnalité du sens d'un syntagme), malgré des avis contraires (Mel'čuk, 1993a; cf. aussi Fónagy, 1982, 1997), ne semble pas jouer pour les pragmatèmes: à côté des pragmatèmes transparents (*Ne quittez-pas, Je vous/t' en prie, Ce sont les risques du métier*) nous avons de nombreux cas de pragmatèmes plus au moins opaques (*Faites chauffer la colle!, Un ange passe*). Une séparation nette entre les pragmatèmes et les phrasèmes est d'autant plus difficile à dégager qu'il y a une gradualité résultant de leur dépendance à des conditions pragmatiques. Où classer, par exemple, *c'est une goutte d'eau qui fait déborder le vase*?

3.6. Schématiquement, notre typologie pourrait se présenter de la façon suivante:



4. LE TRAITEMENT DES EXPRESSIONS IDIOMATIQUES DANS LES DICTIONNAIRES MONOLINGUES DU FRANÇAIS

Comme notre analyse s'efforcera de le montrer plus en détail dans la suite, les expressions idiomatiques lancent un véritable défi à la lexicographie qui, traditionnellement, privilégie la description des «mots»⁴³ isolés, c'est-à-dire des unités monolexicales hors contexte.

4.1. La perspective que nous adopterons dans le présent article est celle d'un locuteur étranger.

4.1.1. Il n'est sans doute pas nécessaire d'insister sur la pertinence de la différence entre un locuteur natif et un locuteur étranger étant donné que l'apprentissage d'une langue maternelle et d'une langue seconde ou étrangère est fondamentalement différent. Par

courir le danger ou *courir le risque* sont peut-être moins transparentes (*izpostavljati se nevarnosti*) que pour un anglophone (*to run a risk*).

42 La réponse de Mel'čuk est «formelle»: si un syntagme peut être formalisé par une FL (syntagmatique), il s'agit d'une collocation (dans les exemples ci-dessus nous avons affaire à $Oper_1$). Sur la question de «l'héritage» lexical dans le cadre théorique *Sens-Texte*, voir Mel'čuk & Wanner, 1996.

43 La notion de *mot* est en lexicographie réduite à une notion «formelle» de mot graphique.

conséquent, les besoins d'un utilisateur étranger diffèrent en plusieurs points des besoins d'un utilisateur natif (cf. Lamy, 1985; Hausmann, 1990; Herbst & Popp éd.s., 1999):

- la langue utilisée dans les définitions et les autres rubriques de l'article dictionnaire devrait être adaptée à la compétence langagière réduite de l'étranger⁴⁴;
- à la différence des dictionnaires pour locuteurs natifs qui devraient fournir plus de mots ou d'acceptions rares et posséder ainsi une macrostructure (voir ci-dessous) plus riche, dans les dictionnaires pour étrangers, une importance particulière devrait être attachée aux informations dans la microstructure (voir ci-dessous): aux connotations, aux registres, aux valeurs stylistiques, à la combinatoire, aux exemples illustrant l'usage réel;
- si les dictionnaires pour locuteurs natifs (et d'ailleurs les dictionnaires monolingues généraux) aident principalement au décodage, les dictionnaires pour étrangers devraient intégrer, en outre, une forte composante d'encodage: un intérêt particulier devrait être porté à la syntagmatique, en premier lieu à toutes sortes de séquences dont les constructions ou les emplois sont imprédictibles;

4.1.2. Les recherches qui évaluent l'utilisation des dictionnaires monolingues (surtout des dictionnaires d'apprentissage) et des dictionnaires bilingues chez les apprenants étrangers (voir Béjoint, 2000, pp. 142-154)⁴⁵ montrent que le bilingue reste un outil préféré, surtout à l'encodage, pour la majorité des apprenants (d'après certaines enquêtes, le taux s'élève à 75%). Il n'est pas surprenant que les monolingues soient davantage utilisés par des apprenants de niveau plus avancé que par ceux de niveaux moyen ou débutant. Béjoint & Moulin (1987, p. 104) ont fait judicieusement remarquer que le bilingue se prête mieux à une consultation rapide (ce qui est d'ailleurs l'utilisation «prototypique» de tout dictionnaire), tandis que les monolingues permettent une meilleure «initiation» de l'utilisateur étranger au système lexical d'une langue qui n'est pas la sienne. Les comparaisons des qualités respectives des deux types de dictionnaires (Atkins, 1985; Béjoint & Moulin, 1987) montrent que, malgré des difficultés auxquelles pourrait se heurter un utilisateur non natif qui consulte les dictionnaires monolingues, ceux-ci, à tout prendre, répondent mieux aux besoins spécifiques du locuteur étranger (surtout avancé) que les dictionnaires bilingues:

- le monolingue offre des informations plus complètes sur la combinatoire, sur les connotations et les associations, sur les valeurs stylistiques, sur les expressions idiomatiques etc., ce qui s'avère très utile surtout pour l'encodage;
- le monolingue «court-circuite» l'anisomorphisme des langues, auquel le bilingue ne peut pas échapper;
- à la limite, le bilingue pourrait même empêcher l'intériorisation de la langue étrangère.

⁴⁴ La tendance à un contrôle rigoureux du vocabulaire définitionnel est moins sensible dans les dictionnaires d'apprentissage français que dans leurs homologues («*learners' dictionaries*») anglo-saxons.

⁴⁵ Ces recherches sont presque exclusivement consacrées aux dictionnaires pour les apprenants de l'anglais, les recherches qui porteraient sur les dictionnaires pour les apprenants du français étant extrêmement rares – de même, d'ailleurs, que les dictionnaires eux-mêmes. Voir plus loin.

4.1.3. Notre analyse portera sur un corpus de dictionnaires monolingues français que pourrait mettre à profit aussi un utilisateur non francophone. Le corpus sera constitué d'un dictionnaire destiné exclusivement aux apprenants étrangers (*Dictionnaire du français* – désormais: *DF*)⁴⁶ et d'un certain nombre de dictionnaires pédagogiques ou généraux qui, à en croire les préfaces, visent aussi les étrangers et qu'utilisent effectivement les apprenants slovènes, en premier chef: *Le Nouveau Petit Robert (NPR)*, le *Micro-Robert (MR)*, *Le Petit Larousse (PL)* et le *Lexis*⁴⁷.

4.1.4. En ce qui concerne l'aspect métalexigraphique de notre analyse, nous prendrons appui principalement sur les notions «classiques» de *macro-* et de *microstructure*, introduites par J. Rey-Debove (1971), ainsi que sur une notion plus récente de *structure d'adressage* (appelée plus brièvement *adressage*), élaborée par F. J. Hausmann et H. E. Wiegand (Hausmann & Wiegand, 1989; Hausmann, 1990a). Rappelons brièvement les trois notions.

La structure du texte dictionnaire, selon J. Rey-Debove (1971), est double: la suite verticalement ordonnée des entrées (*mots-vedettes*, que nous appellerons *lemmes*) forment la macrostructure⁴⁸ d'un dictionnaire; l'ensemble des informations horizontalement ordonnées qui suivent une entrée (= lemme), constitue la microstructure.

Le concept d'adressage sort du cadre rigide de la dichotomie de macro- et microstructure: il décrit la réunion de deux segments du texte dictionnaire en un énoncé lexicographique unique, porteur d'une information pertinente – les deux segments sont en relation *thème-rhème*. Le thème de l'énoncé s'appelle *adresse*. Au cas où l'adresse et le lemme coïncident, on parle d'adressage *lemmatique*. Dans les dictionnaires qui se servent des procédés de regroupement (par exemple, *Dictionnaire du français contemporain*, *Lexis*, *Micro-Robert*), on a affaire à l'adressage *sous-lemmatique*. Dans tous les cas où l'adresse n'appartient plus à la macrostructure, mais fait partie de la microstructure, on parle d'adressage *non lemmatique* (ou de *sous-adressage*): l'adresse de l'adressage non lemmatique s'appelle *sous-adresse*.

4.2. LA DESCRIPTION DES COLLOCATIONS

Le traitement adéquat des collocations est d'autant plus indispensable que les collocations représentent une «écrasante» majorité des expressions idiomatiques (*cf.* Mel'čuk, 1998, p. 31).

4.2.1. Un premier problème que soulèvent les collocations est lié aux critères de choix qui président à l'inclusion des collocations dans un dictionnaire. Idéalement, un dic-

⁴⁶ Le *DF* est d'ailleurs, à notre connaissance, le seul «vrai» dictionnaire pour apprenants du français qui existe actuellement sur le marché français. Pour une petite histoire de ce type de dictionnaires pour le français, voir Hausmann, 1990, p. 1388.

⁴⁷ Le *Micro-Robert* était même officiellement recommandé aux lycéens bavarois.

⁴⁸ Selon J. Rey-Debove (1971, p. 21), la macrostructure est un synonyme de *nomenclature*. Or, nous préférons une terminologie plus rigoureuse qui distingue ces deux termes: la nomenclature est simplement un ensemble non structuré d'unités lexicales admises dans un dictionnaire, la macrostructure étant l'ensemble structuré de mots-entrées ou lemmes (voir Béjoint, 2000, pp. 11-13).

tionnaire pour étrangers devrait inclure une grande quantité de collocations, mais en raison des contraintes de place, le lexicographe est presque toujours obligé de faire des choix. On peut se demander, à juste titre, s'il existe des systèmes plus ou moins rigoureux de critères, auxquels les rédacteurs des dictionnaires actuels obéissent dans leur choix des collocations: d'après l'analyse que nous avons effectuée, il semble qu'il n'en est rien.

4.2.1.1 La métalexigraphie propose normalement deux séries de critères (pour un état de la question, voir Gabrovšek, 1998) dont une seule, la première, est applicable en lexicographie monolingue:

- les critères basés sur les statistiques qui tiennent compte des fréquences d'emploi des collocations⁴⁹;
- les critères qui prennent en ligne de compte des aspects contrastifs propres à une paire donnée de langues: la priorité est donnée aux collocations qui manifestent un plus grand degré d'imprédictibilité dans la traduction (relativement aux propriétés sémantiques, syntaxiques ou pragmatiques).

4.2.1.2. Les FL de Mel'čuk offrent à la lexicographie une troisième possibilité, absente de la plupart des études métalexigraphiques, à savoir des critères d'ordre sémantique. Nous avons déjà dit que les FL constituent un nombre limité d'invariants syntactico-sémantiques (très abstraits et généraux) qui formalisent la cooccurrence lexicale restreinte⁵⁰. Le lexicographe dispose ainsi d'un outil qui lui permet de sélectionner les collocations selon les sens qu'elles expriment. Les avantages d'une telle approche sont loin d'être négligeables, principalement pour un locuteur étranger en situation d'encodage. En raison de l'économie de place à laquelle tout projet lexicographique (commercial) est contraint, le lexicographe peut décider de ne garder qu'une collocation pour une FL donnée et, de ce fait, inclure plus de FL différentes. Par exemple, pourquoi ne pas éliminer dans l'article *colère* du MR une ou deux collocations pour la FL Sing: *accès, crise, mouvement de colère*, et introduire, à la place, des collocations qui expriment d'autres FL, tout aussi importantes: *rentrer* (NonPerm₁Manif), *retenir* (NonPerm₁Fact₀), *provoquer/susciter/soulever/exciter* (CausFunc₁), *étouffer* (LiquFunc₀) *la colère*; *la colère s'empare* [de qqn]/*envahit* [qqn]/*saisit* [qqn] (IncepFunc₁), *s'apaise/se calme* (IncepPredMinus). Quant au choix des collocations à éliminer ou à garder, les critères sémantiques pourraient et devraient se combiner avec d'autres critères: dans un dic-

⁴⁹ La fréquence est déterminée soit «intuitivement» (la compétence du rédacteur) soit à l'aide de corpus électroniques. Si l'utilisation de vastes corpus électroniques est très répandue, voire généralisée, en lexicographie anglo-saxonne, les rédacteurs français, à l'exception notable de l'équipe du *Trésor de la langue française*, semblent plus prudents: leurs sources principales demeurent leur compétence ainsi que la compétence des informateurs. La citation suivante, tirée de la préface au DF de J. Rey-Debove, est significative de ce point de vue: «Les exemples ne viennent pas directement d'un corpus, mais de la réflexion du rédacteur pour produire une phrase attendue avec le mot en question». (DF, p. XI)

⁵⁰ Il ne faut pas perdre de vue, et Mel'čuk l'a d'ailleurs souligné à plusieurs reprises, que les FL représentent un outil lexical et non pas des «entités» sémantiques: elles ne sont pas à même de rendre compte de toutes les nuances sémantiques qu'impliquent les relations syntagmatiques. Voir Mel'čuk, 1996, pp. 79-86.

tionnaire pour apprenants étrangers, on aura sans doute tendance à retenir des collocations «neutres»⁵¹ et fréquentes. Cette approche, cependant, présenterait des inconvénients pour un utilisateur qui voudrait, par précaution, vérifier la collocation éliminée dans un dictionnaire mais qui n’y trouverait qu’une des collocations synonymes.

4.2.2. L’ADRESSAGE.

4.2.2.1. Bien que la linguistique accorde aux collocations le statut d’unités de langue, au même titre qu’aux «mots», elles ne font jamais, à notre connaissance, partie intégrante de la macrostructure – l’adressage des collocations est uniquement non lemmatique⁵²: elles sont donc insérées dans les microstructures des articles. Mais puisque la collocation est une unité polylexicale, la question qui se pose aussitôt est de savoir sous quel élément la traiter.

4.2.2.2. La question du placement⁵³ paraît d’autant plus cruciale quand on prend en considération l’asymétrie sémantique des deux constituants (que nous avons déjà évoquée ci-dessus): Hausmann (1989) insiste avec vigueur sur le fait que les fonctions lexicographiques de la collocation diffèrent suivant que la collocation est décrite dans l’article de la *base* ou dans l’article du *collocatif*. Le traitement dans l’article du collocatif (Hausmann parle de la perspective *sémasiologique*) est surtout utile au décodage parce qu’il sert à la discrimination des différents sens du collocatif. Les dictionnaires qui fondent l’analyse sémantique sur des propriétés distributionnelles des unités lexicales illustrent bien cette fonction: un bon exemple nous est fourni par les dictionnaires rédigés sous la direction de Jean Dubois (*DFC, Lexis*):

AGITER [...] v. tr. (lat. *agitare*, fréquentatif de *agere*, agir; v. 1200) 1. (sujet nom de chose ou de pers.) *Agiter quelque chose (un objet)*, le remuer vivement en tous sens: *Il agitait ses feuilles pour en faire sécher l’encre* (Camus) [syn. SECOUER]. *Les enfants agitaient de petits drapeaux au passage du président.* [syn. BRANDIR]. [...] – 2. (sujet nom de pers.) *Agiter une menace*, présenter quelque chose comme un danger imminent: *Il agita la menace de sa démission.* – 3. (v. 1600). *Agiter une question, un problème*, etc., les discuter avec d’autres personnes: *On a agité la question de savoir s’il devait poser sa candidature* (syn. SOULEVER). – 4. (sujet nom de chose) *Agiter quelqu’un*, lui causer une vive inquiétude, une émotion ou une peine profonde; l’exciter facilement: *Je fus agité tout entier par la curiosité douloureuse de savoir ce qu’elle aurait pu faire* (Proust) [syn. BOULEVERSER, PRÉOCCUPER, REMUER]. *Les souvenirs l’agitaient, le poursuivaient* [...] (*Lexis*)

51 Nous pensons ici aux collocations qui ne sont marquées ni quant au niveau de langue, ni quant à l’aspect diachronique, ni quant au domaine, ni quant au style.

52 Les collocations ne figurent pas non plus dans les macrostructures des dictionnaires spécialisés, comme le *DEFC* (qui traite pourtant macrostructurellement les phrasèmes et les pragmatèmes) ou le *BBI Dictionary of English Word Combinations*.

53 Cette question ne se pose réellement que dans le cas des dictionnaires «traditionnels», sur support papier, où la macrostructure fournit souvent le seul accès possible aux informations de microstructure.

Chaque définition est directement précédée d'une rubrique où le mot-lemme apparaît dans un contexte minimal⁵⁴: dans la première et la dernière division de sens, celle-ci indique, au niveau métalinguistique, les contraintes qui régissent le choix de l'objet (*agiter quelque chose*: objet = [- humain], *agiter quelqu'un*: objet = [+ humain]), tandis que dans les deux autres, les contextes consistent en une base (*agiter une menace*) et en un paradigme ouvert de bases possibles (*agiter une question, un problème*).

L'inclusion de la collocation dans l'article de la base (perspective *onomasiologique*, selon Hausmann), en revanche, favorise l'encodage: la production d'une collocation se fait dans le sens *base* → *collocatif* – nous avons déjà dit que le locuteur en situation d'encodage choisit d'abord la base, puis, en fonction de la base choisie, cherche le collocatif approprié.

4.2.2.3. Les différents statuts de la base et du collocatif quant à l'autonomie sémantique se reflètent inévitablement dans la description dictionnaire: en règle générale, la collocation est définie ou glosée⁵⁵ beaucoup plus souvent si elle est placée dans l'article du collocatif – dans l'article de la base, elle n'est que mentionnée. Voici quelques exemples des deux modalités de traitement tirés du *NPR*:

«*Rompre le silence, le faire cesser par un son, et spécialt par la parole.*» [s.v. *rompre*]

«*Rompre le silence.*» [s.v. *silence*]

«*Dormir profondément, intensément (cf. À poings* fermés).*» [s.v. *profondément*]

«*Dormir profondément, d'un profond sommeil.*» [s.v. *dormir*]

«*Gratter une allumette, l'enflammer.*» [s.v. *gratter*]

«*Gratter une allumette.*» [s.v. *allumette*]

Notre constatation est corroborée par les résultats d'une analyse quantitative que nous avons effectuée et qui a porté sur le traitement dictionnaire d'un corpus de 125 collocations⁵⁶ dans trois dictionnaires monolingues (*NPR*, *DF* et *MR*; voir aussi plus loin). Le tableau suivant donne le nombre des collocations glosées relativement à leur placement:

	DF	MR	NPR
BASE	5	1	6
COLLOCATIF	17	12	8
TOTAL	22	13	4

Nous pouvons observer que dans le *DF*, par exemple, sur un total de 22 collocations glosées, 17 collocations figurent dans l'article du collocatif et 5 seulement dans l'article de la base.

⁵⁴ Les (méta)lexicographes ne font pas de différence entre cette rubrique qui précède la définition et les exemples qui se placent à droite de la définition. Nous proposons, pour distinguer les deux rubriques, d'appeler la première *indicateur contextuel*.

⁵⁵ Une *glose* diffère d'une *définition*, en ce qu'elle n'offre qu'une paraphrase partielle, plus sommaire de l'adresse. Voir Hausmann & Wiegand, 1989.

⁵⁶ Nous estimons qu'un tel échantillon peut prétendre à une certaine représentativité.

À défaut de gloses (ou de définitions), la fonction d'explication sémantique peut être remplie, au moins en partie, également par des renvois analogiques⁵⁷ dont l'emploi est très étendu dans les dictionnaires examinés. La majorité des renvois adressés aux collocations peut être répartie en deux types:

- les uns proposent des quasi-synonymes d'une collocation donnée:
 - «*Lancer, jeter un défi.* ⇒ défier, provoquer» (*NPR*, s.v. *défi*)
 - «*Déposer une plainte contre un agresseur, contre X.* ⇒ accuser, dénoncer» (*NPR*, s.v. *plainte*)
- les autres signalent des variantes paradigmatiques d'un des éléments de la collocation (normalement, c'est le collocatif):
 - «*La difficulté réside en ceci.* ⇒ consister, se situer» (*NPR*, s.v. *résider*)
 - «*Commettre un attentat, un crime, un meurtre.* ⇒ consommer, perpétrer.» (*NPR*, s.v. *commettre*)

Notre analyse a montré que les renvois étaient nettement plus nombreux dans les articles des collocatifs que dans ceux des bases: dans le *NPR*, le rapport est de 13 contre 29, dans le *MR* de 5 contre 23, dans le *DF* même de 5 contre 29.

4.2.2.4. Revenons à notre question de départ qu'il convient de compléter de façon suivante: où traiter des collocations dans un dictionnaire pour usagers (apprenants) étrangers?

4.2.2.4.1. En ce qui concerne les dictionnaires monolingues généraux, les analyses métalexigraphiques montrent que la perspective dominante est sémasiologique: les collocations sont décrites essentiellement sous les collocatifs. Avant de se lancer dans une discussion théorique, examinons brièvement les résultats de notre analyse (voir aussi ci-dessus) qui avait pour but de déterminer statistiquement où les trois dictionnaires (*NPR*, *MR* et *DF*), qui devraient s'adresser (aussi) aux utilisateurs non francophones, insèrent les collocations. Le tableau suivant donne les résultats (en chiffres et en pourcentages arrondis) pour 125 collocations examinées⁵⁸ quant à la place de leur traitement (soit dans l'article de la base, soit dans l'article du collocatif, soit dans les deux articles):

	DF	MR	NPR
BASE	11 (9%)	15 (12%)	19 (15%)
COLLOCATIF	40 (32%)	63 (50%)	72 (58%)
BASE + COLLOCATIF	74 (59%)	47 (38%)	34 (27%)

Dans le *NPR*, un tiers de collocations (32%) seulement est décrit uniquement dans l'optique sémasiologique, tandis qu'on observe une nette tendance (59%) à combiner les deux perspectives (sémasiologique et onomasiologique) et à traiter les collocations sous tous les éléments. La perspective sémasiologique est néanmoins plus manifeste

⁵⁷ Une discussion, même superficielle, de la question des renvois analogiques, et de l'analogie en général, nous éloignerait trop du sujet. Nous nous contenterons d'un simple «renvoi» à Rey-Debove, 1989.

⁵⁸ La liste des collocations qui constitue le corpus est donnée en appendice à la fin de l'article.

dans les deux autres dictionnaires (50% pour le *MR* et 58% pour le *DF*). En raison des dimensions plus modestes de ces deux dictionnaires, sans doute, les exemples du «double» traitement (la description sous tous les éléments) y sont plus rares (38% et 27% respectivement). En s'appuyant sur l'observation que le nombre de collocations mentionnées à l'article du collocatif est plus élevé dans les dictionnaires plus petits, où le facteur espace joue un rôle plus décisif, nous pouvons avancer l'hypothèse que les dictionnaires monolingues, même ceux qui sont destinés aux apprenants étrangers, privilégient toutefois la perspective sémasiologique: le rédacteur qui ne peut pas se payer le luxe d'un double traitement et doit choisir entre la mention de la collocation sous la base et la mention de la collocation sous le collocatif préférera plus souvent la dernière possibilité – la perspective sémasiologique est, pour ainsi dire, une perspective «par défaut».

Sans prétendre à l'exhaustivité, ajoutons deux autres facteurs qui, à notre avis, influent sur le placement des collocations dans les trois dictionnaires examinés:

- la longueur de l'article: la collocation n'est qu'exceptionnellement insérée dans l'article qui risque d'être trop long (*faire, passer, prendre...*): c'est peut-être une des raisons pourquoi le *DF* traite la collocation *faire une liste* sous la base (*liste*) et son synonyme *dresser une liste* sous le collocatif (*dresser*).
- l'opacité sémantique: les collocations plus ou moins opaques sont presque toujours traitées sous le collocatif (*reproches sanglants, passion dévorante, assouvir la curiosité, la vérité jaillit*).

4.2.2.4.2. Ces résultats vont pourtant à l'encontre des points de vue qui dominent en théorie. Il semble, en effet, que les rédacteurs des dictionnaires ne se rendent pas compte des besoins spécifiques des apprenants étrangers dans le domaine des collocations.

La position d'un des plus grands spécialistes dans la matière, de F. J. Hausmann (1990), sur le placement des collocations est nette: les collocations ne posent pas de problèmes au décodage, étant donné qu'elles sont sémantiquement transparentes, cependant, elles sont «idiomatiques» à l'encodage et doivent, par conséquent, être décrites dans les articles des bases – l'utilisateur qui veut trouver une collocation la cherchera forcément dans l'article de la base.

Le point de vue de Hausmann devrait être, à notre avis, quelque peu nuancé. Tout en reconnaissant que la plupart des collocations se révèlent en effet transparentes au décodage, il ne faut pas faire abstraction d'un nombre non négligeable de collocations qui manifestent un certain degré d'opacité sémantique (*reproches sanglants, activité fiévreuse, injustice criante, le malheur frappe, établir un impôt, courir un danger, surmonter la peur*): pour un apprenant étranger, le traitement adéquat de ces collocations dans l'article du collocatif n'est sans doute pas dépourvu d'intérêt, puisqu'il est fort probable⁵⁹ qu'il cherchera sous le collocatif. Bogaards (1999) dans une critique de la position (trop catégorique) de Hausmann⁶⁰, soutient, que, même pour l'encodage, il est

⁵⁹ L'élément qui en raison de son opacité pose problème est presque toujours le collocatif.

⁶⁰ Pour la réponse «catégorique» de Hausmann, voir Hausmann, 1999, p. 210.

utile d'insérer les collocations dans l'article du collocatif. Par exemple, un Slovène qui voudrait dire *kositi travo* en français (*tondre le gazon*) ne trouverait pas la collocation s'il la cherchait sous l'*herbe*, qui s'impose pourtant comme un équivalent «évident» du mot *trava*. De même, il est difficile (sinon impossible) de prévoir où un utilisateur, qui se sert du dictionnaire monolingue pour la traduction vers une langue étrangère, cherchera les collocations. En tout cas, à l'état actuel des études métalexigraphie sur l'utilisation des dictionnaires, nous savons encore trop peu sur les stratégies de recherche que différents types d'utilisateurs mettent en œuvre dans différentes situations d'utilisation.

4.2.2.4.3. La conclusion qui découle tout naturellement de la discussion précédente ne doit surprendre personne: dans un dictionnaire pour apprenants étrangers, les collocations devraient être traitées, à la fois, dans l'article de la base et dans l'article du collocatif. Nous admettons que ce «double traitement» prend trop de place et qu'il sera par conséquent difficilement applicable en lexicographie commerciale, mais une solution, pourtant, peut être suggérée. Le rédacteur qui décidera d'inclure une collocation sous la base, peut recourir au système de renvois, et renvoyer de la base au collocatif, ou inversement. Ce dispositif permettrait, en outre, un traitement non seulement plus économique mais en même temps plus systématique de toutes les collocations dont les bases partagent les mêmes collocatifs⁶¹: par exemple, *éprouver* peut se combiner avec presque tous les substantifs qui dénotent des sentiments, *commettre* sert de collocatif à un vaste ensemble de bases (*maladresse, injustice, erreur, faute, attentat, crime...*), de même *essuyer* (dans le sens de «subir») ou *net* («sans équivoque») etc. Les exemples sont très nombreux. Il suffirait de renvoyer des bases à l'article du collocatif où toutes les bases possibles devraient être signalées soit sous forme d'une liste plus ou moins exhaustive (par exemple, *net* – sans équivoque /*refus, position, réponse*/) soit ces bases seraient remplacées par un ou plusieurs termes génériques (*éprouver* – avoir /un sentiment, une sensation/).

4.2.2.5. À l'intérieur d'un article, les collocations sont placées dans des divisions sémantiques appropriées et ne sont jamais regroupées dans une division à part, qui se présenterait comme une sorte d'annexe, placée en fin d'article. Dans la terminologie de Wiegand (1990), nous pouvons dire que les collocations sont *intégrées* dans la micro-structure. Malgré un certain degré d'opacité des collocations, le rattachement d'une collocation à un sens du mot sous lequel elle est décrite, normalement, ne pose pas de problèmes. L'intégration entraîne pourtant une conséquence fâcheuse: les collocations sont mises sur le même pied que les exemples de combinaisons libres et très souvent aucun dispositif (par exemple, différents moyens typographiques) n'est employé pour les séparer (pour une discussion sur les formes de présentation des collocations, voir plus loin). Voici un exemple du *DF*:

EXCUSE [...] n.f. UNE EXCUSE 1. Regret que l'on témoigne à qqn de l'avoir offensé, gêné, contrarié. *J'exige des excuses. Il veut que tu lui fasses des excuses. Je*

⁶¹ Ces bases appartiennent normalement aux mêmes champs sémantiques. Voir aussi la note 42.

vous fais toutes mes excuses. Veuillez accepter mes excuses. Il m'a présenté ses excuses. Elle s'est confondue en excuses. Il a refusé mes excuses.[...] (s. v. excuse)

4.2.2.6. Une autre question qui mérite réflexion, mais à laquelle les lexicographes prêtent peu d'attention, est la question de l'ordonnement des collocations à l'intérieur d'une division sémantique. Dans les quatre dictionnaires examinés (*NPR*, *MR*, *DF*, *Lexis*), il nous a été impossible d'isoler des critères systématiques qui présideraient à l'arrangement des collocations les unes par rapport aux autres: les collocations sont ordonnées sans ordre apparent et souvent mêlées aux combinaisons libres ou même aux autres types d'expressions idiomatiques.

Dans un dictionnaire pour apprenants étrangers, la meilleure façon d'ordonner les collocations serait, à notre avis, l'arrangement qui obéirait aux critères syntactico-sémantiques: le lexicographe grouperait d'abord les collocations selon les constructions syntaxiques (par exemple, mentionner d'abord les collocations verbe + nom, ensuite nom + verbe, ensuite non + préposition + nom...), après quoi, il procéderait aux regroupements des collocations synonymes. Nous estimons que cet ordonnancement est préférable à l'ordonnement alphabétique, pourtant plus facile à consulter, mais qui entraînerait la «dispersion» des collocations synonymes et affaiblirait, par conséquent, considérablement la composante d'encodage du dictionnaire.

4.2.3. FORMES DE PRÉSENTATION DES COLLOCATIONS

Les collocations peuvent être présentées soit sous une forme neutralisée⁶² (*infliger une amende*) soit à l'intérieur d'une phrase complète (*Le gendarme a infligé une amende à l'automobiliste*). La première façon de présenter les collocations a l'avantage de souligner leur statut d'unités de langue. Cependant, un certain nombre de collocations ne se prête pas facilement à la neutralisation: comment neutraliser, par exemple, les collocations nom + verbe, comme *le malheur frappe*, *le silence règne*, *la difficulté réside*? Pour un utilisateur étranger, en outre, cette forme de présentation comporte un sérieux inconvénient: elle n'offre aucun renseignement sur l'emploi réel (en contexte) de la collocation. Il ne faut pas oublier non plus, que l'apprenant peut avoir de la difficulté à trouver le sujet approprié (animé/non animé, humain/non humain), la forme du verbe, le déterminant qui convient etc. Cette forme de présentation est dominante dans le *NPR* et le *MR*. La typographie dans la grande majorité des cas ne diffère pas de celle des exemples de combinaisons libres: un certain nombre de collocations figure pourtant en petites capitales que le *NPR* et le *MR* réservent normalement aux phrasèmes ou aux constructions syntaxiques importantes.

La seconde façon de présenter les collocations, c'est-à-dire la présentation sous forme de phrases complètes, par contre, ne permet pas de distinguer les collocations

⁶² Le terme de *neutralisation* est emprunté à J. Rey-Debove (1971, pp. 303 *ssq.*): il désigne un procédé métalinguistique qui à partir de plusieurs énoncés dégage un modèle commun. Dans le cas des collocations, on pourrait parler aussi de *lemmatisation*.

des combinaisons libres (voir plus haut) ni de signaler clairement à l'utilisateur quel segment de la phrase relève du figement, ou plus précisément de la cooccurrence lexicale restreinte. Un autre inconvénient de cette présentation est de prendre trop de place et d'être donc très «coûteuse». Cette présentation domine, par exemple, dans le *DF*. Cependant, ce dictionnaire propose aussi un moyen de présentation des collocations qui paraît très approprié aux besoins spécifiques de l'apprenant étranger puisqu'il réunit la plupart des avantages des deux formes évoquées. La collocation qui est donnée dans une phrase complète est marquée typographiquement – elle est en petites capitales:

«*FAITES LA COMPARAISON de ce tableau avec celui-là.*» (s.v. *comparaison*)

«*Mon fils PASSE UN EXAMEN dans trois jours.*» (s.v. *examen*)

«*Il a demandé une augmentation à son patron et il S'EST HEURTÉ À UN REFUS.*»
(s.v. *refus*)

«*Le départ du directeur POSE UN PROBLÈME*» (s.v. *problème*)

Malheureusement, ce dispositif n'est pas employé de façon systématique: une même collocation, par exemple, peut être présentée sous des formes différentes:

«*Un chanteur entre en scène sous un tonnerre d'applaudissements.*» (s.v. *tonnerre*)

«*Son discours a SOULEVÉ UN TONNERRE D'APPLAUDISSEMENTS.*» (s.v. *applau-disse-ment*)

Outre cela, les petites capitales ne sont pas réservées exclusivement aux collocations, mais sont utilisées aussi pour mettre en évidence (de même que dans le *NPR* et le *MR*) les phrasèmes et les constructions typiques⁶³:

«*Le voilier navigue à L'ESTIME.*» (s.v. *estime*)

«*Il va passer par la CHIRURGIE ESTHÉTIQUE pour se faire refaire le nez*» (s.v. *esthétique*)

«*Elle se met du VERNIS À ONGLES. Où sont les CISEAUX À ONGLES?*» (s.v. *ongle*)

«*La mairie PRÉVOIT DE construire de nouveaux logements*» (s.v. *prévoir*)

«*Il PROJETTE DE l'épouser.*» (s.v. *projeter*)

4.4. LA DESCRIPTION DES PRAGMATÈMES

Le fait que les pragmatèmes sont écartés d'un grand nombre de typologies des expressions idiomatiques, entraîne inévitablement l'absence d'un statut lexicographique spécifique.

4.4.1. La presque totalité des pragmatèmes est traitée au niveau de la microstructure où aucun indice (une typographique spécifique ou un indicateur métalinguistique) ne permet aux utilisateurs de distinguer les pragmatèmes des autres types d'expressions idiomatiques ni, souvent, des exemples de combinaisons libres. Il est significatif que

⁶³ Pour une proposition de normalisation voir § 4.5.3.

même dans les rares cas, où les pragmatèmes sont insérés dans la macrostructure⁶⁴, le lemme ne soit suivi que d'une indication de catégorie grammaticale:

ALLO ou **ALLÔ** [...] interj. • 1880; **angl. amér.** *hallo, hello*, onomat. ♦ Interjection conventionnelle servant d'appel dans les communications téléphoniques. *Allo, allo, ne coupez pas.* [...] (*NPR*)

BONSOIR [...] n. m. • fin XV^e; de l. *bon* et *soir* ♦ Salutation adressée à qqn qu'on rencontre en fin d'après-midi ou le soir (de préférence à *bonjour**), ou pour quitter qqn en fin de journée (de préférence à *au revoir**, *adieu**) [...] (*NPR*)

4.4.2. Comme nous l'avons déjà dit, la plupart des pragmatèmes sont intégrés dans les microstructures des articles d'un de leurs constituants: en règle générale, ils sont rattachés à un sens du constituant et, par conséquent, décrits à l'intérieur des divisions sémantiques.

4.4.2.1. Les modalités du traitement dictionnaire des pragmatèmes dépendent fortement de leurs propriétés linguistiques. Les pragmatèmes ne figurent que rarement sous forme neutralisée étant donné qu'ils n'admettent presque pas de modifications ou de transformations – ils sont employés «tels quels» dans des situations de communication données, et jouissent, par ailleurs, d'une grande autonomie syntaxique. Quel intérêt aurait-on à neutraliser *Tu m'en diras des nouvelles?* si une seule variation est acceptable (*Vous m'en direz des nouvelles?*)? Ou bien *Tu parles!* et *Un ange passe*, par exemple, qui n'admettent aucune modification? Même si le dictionnaire donne la forme neutralisée du pragmatème, il est nécessaire de fournir, en plus, au moins un exemple de la forme sous laquelle le pragmatème apparaît normalement en discours. Prenons comme exemple un extrait du *NPR*:

«*Numéroter ses abattis*: se préparer à une lutte comme si on risquait de perdre l'intégrité et la disposition de ses membres. *Tu peux numéroter tes abattis!* (menace).» (s. v. *abattis*)

La forme neutralisée et la définition qui suit sont trop générales, et ne donnent pas d'information complète ni sur le sens ni sur l'emploi de ce pragmatème, qui apparaît le plus fréquemment à la 2^e personne du singulier et qui exprime une menace: l'exemple est donc indispensable.

4.4.2.2. L'information que donne un dictionnaire sur un pragmatème donné doit, en outre, inclure le renseignement sur la situation qui «déclenche» son énonciation. Prenons comme exemple deux extraits où la définition du pragmatème est réduite à une simple paraphrase sémantique:

⁶⁴ Ce sont exclusivement les pragmatèmes constitués d'un seul mot (*bonjour, salut, tchao, adieu*). La situation est différente dans le *DECFC* où tous les pragmatèmes ainsi que tous les phrasèmes sont adressés lemmatiquement.

«*Je me comprends*: je sais ce que je veux dire.» (NPR, s. v. *comprendre*)

«*Vous m'en direz des nouvelles*: vous m'en direz sûrement du bien, vous m'en ferez compliment.» (NPR, s. v. *nouvelle*)

Les précisions sur la situation où les pragmatèmes sont employés sont utiles surtout pour les besoins de l'encodage, mais également pour les besoins de la traduction vers la langue maternelle où l'information sur la valeur réelle (c'est-à-dire, en discours) du pragmatème est souvent nécessaire pour le choix adéquat de l'équivalent. Voici deux exemples où figure l'information sur la situation d'énonciation:

«*Ne vous gênez pas!* se dit à qqn d'indiscret, d'inconvenant, qui en prend un peu trop à son aise.» (NPR, s. v. *gêner*)

«*Cela ne se passera pas comme ça, vous AUREZ DE MES NOUVELLES!* menace qui avertit que l'on va répondre à l'attaque de qqn.» (DF, s. v. *nouvelle*)

Ces informations peuvent être ramenées à une sorte de commentaire métalinguistique:

«*Je te prie, je vous prie, je vous en prie*, formules de politesse.»

(NPR, s. v. *prier*)

«*C'est pour aujourd'hui (ou pour demain)?* exprime l'impatience.»

(NPR, s. v. *aujourd'hui*)

Le commentaire peut prendre même la forme d'une «étiquette»:

«(Menace) *Vous aurez, vous entendrez de mes nouvelles! Soyez tranquilles, il aura bientôt de mes nouvelles!*» (NPR, s. v. *nouvelle*)

«Spécialt (formules de politesse) *Veillez agréer mes salutations distinguées, l'expression de mes sentiments respectueux.*» (NPR, s. v. *agréer*)

«(souhait) *Bonne fête! Bon voyage!, Bonne année et bonne santé*» (DF, g. *bon, bonne*)

Mais, très souvent, même cette «étiquette» fait défaut: dans le cas où aucune information n'est adressée au pragmatème et où le pragmatème n'est pas mis en relief par des moyens typographiques, il y a lieu de se demander si les dictionnaires ne considèrent pas les phrasèmes simplement comme des exemples qui illustrent le sens ou l'emploi d'un des constituants du pragmatème:

«*Comment allez-vous?; fam. comment ça va?; pop. comment que ça va?*» (NPR, s. v. *comment*)

«*Ne vous dérangez pas pour moi.*» (NPR, s. v. *déranger*)

«*Amicalement vôtre*» (NPR, s. v. *vôtre*)

«2. Demander, inviter. «*Puis-je entrer? – Je vous en prie.*», «*Excusez-moi. – Je vous en prie, ce n'est rien*»» (DF, s. v. *prier*)

«*Veillez agréer, monsieur, mes salutations distinguées*» (DF, s. v. *vouloir*)

4.4.2.3. Une dernière question que nous nous proposons d'étudier est liée à l'accessibilité des pragmatèmes dans les dictionnaires examinés. Une remarque s'impose aussitôt. Le traitement à l'intérieur de l'article d'un des constituants du pragmatème privilégie exclusivement le décodage: l'utilisateur étranger qui se sert du dictionnaire monolingue pour l'encodage et qui cherche un pragmatème approprié à une situation donnée est voué à l'échec, sauf s'il connaît au moins un des constituants (de préférence, l'ensemble des constituants) du pragmatème cherché. Le seul accès possible aux formules finales des lettres, par exemple, passe par les lemmes *agréer, vouloir, prier, sentiment, salutation* etc. Dans la recherche des «mots-clé», très souvent, sa langue maternelle ne lui est d'aucun secours – prenons à titre d'exemple, les pragmatèmes qui jalonnent une conversation téléphonique en français et en slovène, pour voir à quel point les divergences entre deux langues peuvent être grandes:

<i>C'est de la part de qui?</i>	<i>Kdo (pa) kliče? S kom govorim?</i>
<i>Allo, je suis (bien) chez M. X?</i>	<i>Oprosite, ali govorim z g. X-em?</i>
<i>Pourriez-vous me passer M. X?</i>	<i>Ali bi lahko govoril z g. X-em?</i>
<i>Ne quittez pas!</i>	<i>Trenutek prosim!</i>

Pour l'encodage, il serait, à notre avis, préférable de traiter les pragmatèmes séparément, hors de la macrostructure centrale qui obéit forcément au principe sémasiologique (notamment, au classement alphabétique). Les pragmatèmes pourraient être regroupés soit dans des intertextes (par exemple, insérés sous forme d'encadrés dans la macrostructure centrale) soit dans un ou plusieurs posttextes en fin d'ouvrage. Leur arrangement à l'intérieur de ces paratextes, devrait être onomasiologique: les pragmatèmes seraient classés soit selon les situations typiques qui les déclenchent (conversation téléphonique, correspondance, invitation, visite, courses...) soit selon les concepts qu'ils expriment (accord/désaccord, menace, surprise, incrédulité, politesse/impolitesse...). Un dictionnaire destiné aux étrangers devrait impérativement fournir les définitions ou les gloses des pragmatèmes traités ainsi que les informations concernant les niveaux de langue, les valeurs stylistiques et les connotations.

4.5. LA DESCRIPTION DES PHRASÈMES

Les phrasèmes occupent une place privilégiée dans les études phraséologiques – souvent, ils sont même considérés comme des expressions idiomatiques par excellence.

4.5.1. Étant donné que les lexicographes s'accordent à les considérer comme des unités de langue à part entière, il est surprenant de constater que les phrasèmes sont rarement admis dans la macrostructure⁶⁵. Une des raisons réside sans doute dans l'économie de place: l'adressage lemmatique requiert une typographie marquée du lemme (le lemme

⁶⁵ Ceci est vrai aussi pour les dictionnaires spécialisés des phrasèmes (par exemple, *Dictionnaire des expressions et locutions*, éd. A. Rey/S. Chantreau, Le Robert, Paris) qui préfèrent les traiter sous un de leurs constituants. Pour le traitement dans le *DECFC* voir n. 64.

est normalement imprimé en caractères gras et même en capitales grasses) ainsi qu'une indication systématique de toutes les informations qui s'adressent au lemme selon un programme de microstructure préétabli (prononciation, catégorie grammaticale, étymologie...), ce qui prendraient bien trop de place. À cela s'ajoutent les difficultés que soulève souvent la lemmatisation des phrasèmes: par exemple, que faire des phrasèmes qui admettent des substitutions paradigmatiques de leurs éléments (*ceindre/coiffer/porter la tiare, porter les armes/la couronne/la robe/la livrée/la soutane/le froc*)?

Un certain nombre de phrasèmes fait pourtant partie des macrostructures centrales des dictionnaires examinés.

Un premier groupe de ces phrasèmes, que nous pouvons discerner, est constitué des mots composés formés de termes soudés ou de termes reliés par un trait d'union (*malheur, gendarme, deux-roux, arrière-grands-parents*). En ce qui concerne les mots composés dont les éléments sont séparés par des blancs, leur insertion dans la macrostructure est plutôt aléatoire. Le tableau suivant montre les différents types d'adressages qu'adoptent les rédacteurs de cinq dictionnaires (*NPR, MR, PL, DF* et *Lexis*) pour la description d'un petit échantillon de noms composés:

	point de vue	chemin de fer	pomme de terre
NPR	ad. lemmatique	ad. lemmatique	sous-adressage
MR	ad. lemmatique	ad. sous-lemmatique	ad. lemmatique
DF	sous-adressage (avec mention du mot composé dans la macrostructure et renvoi à la microstructure sous un des éléments)	sous-adressage (avec mention du mot composé dans la macrostructure et renvoi à la microstructure sous un des éléments)	sous-adressage
PL	ad. lemmatique	ad. lemmatique	ad. lemmatique
Lexis	ad. lemmatique	ad. lemmatique	ad. lemmatique

Un deuxième groupe est composé des phrasèmes qui contiennent des archaïsmes lexicaux qui, sauf dans le phrasème en question, ne sont plus en usage en français moderne. Ce sont, par exemple, *à la bonne franquette, tout à trac, tout de go, de guingois, en un tournemain, courir le guilledou, peu ou prou* etc.

4.5.2. La majeure partie de phrasèmes ne fait pas l'objet d'articles autonomes, mais est insérée dans les microstructures des articles des constituants.

4.5.2.1. Le premier problème qui se pose pour un utilisateur est de savoir sous quel élément sont traités les phrasèmes. Une analyse détaillée de cette question mériterait sans doute une étude à part, mais un certain nombre de facteurs qui influent sur le placement des phrasèmes dans les dictionnaires analysés peuvent toutefois être inventoriés. Ces facteurs interviennent, d'ailleurs, de façon séparée ou en combinaison:

- démarche «formelle»: le phrasème est enregistré à l'article du premier élément, sauf si c'est un mot grammatical ou si cet élément admet des modifications paradigmatiques. Par exemple, *à tout rompre* sera ainsi traité sous *rompre, ceindre/coiffer/porter la tiare* sous *tiare*;

- longueur d'article: les phrasèmes sont rarement traités sous des éléments qui donnent lieu à des articles (trop) longs, comme *avoir, faire, prendre, mettre...*;
- catégorie grammaticale du lemme: il semblerait que les phrasèmes soient plus souvent traités sous les substantifs que sous les verbes⁶⁶;
- facteur sémantique: dans les dictionnaires qui optent pour l'intégration des phrasèmes, ceux-ci sont insérés dans l'article qui permet une meilleure adaptation à une des divisions de sens (voir ci-dessous).

4.5.2.2. La place du phrasème à l'intérieur d'un article peut être plus ou moins prévisible selon que les phrasèmes font partie des divisions sémantiques (= *intégration*) ou qu'ils sont regroupés dans une sorte d'annexe et nettement détachés du reste de l'article (= *annexion*)⁶⁷. Entre ces deux modalités qui se trouvent aux antipodes, s'échelonne tout une gamme de différentes dispositions des phrasèmes.

4.5.2.2.1. *Intégration.* Cette forme de répartition est largement dominante dans tous les dictionnaires édités par la maison d'édition Le Robert, y compris dans le *DF*. Le phrasème est rattaché à l'un des sens de l'élément sous lequel il est inséré et traité dans une division de sens approprié. Sur le plan fonctionnel, les phrasèmes sont ainsi mis au même niveau que les exemples: ils illustrent les sens ou les emplois des mots-lemmes. La confusion entre ces deux rubriques est d'autant plus grande que les dictionnaires (surtout le *NPR* et le *MR*), normalement, ne recourent pas aux moyens typographiques pour les différencier, les deux rubriques étant, sans distinction, en italique. En effet, rares sont les cas où les phrasèmes ou des parties des phrasèmes sont mis en relief par l'emploi des petites capitales: *CASSER LA TÊTE* (*NPR*, s.v. *casser*), *SE METTRE MARTEL EN TÊTE* (*NPR*, s.v. *martel*), *FAIRE BANDE À PART* (*NPR*, s.v. *bande*), *faire un PONT D'OR à qqn.* (*NPR*, s.v. *pont*), *tomber DANS LES POMMES* (*NPR*, s.v. *pomme*). L'utilisation des petites capitales est plus systématique dans le *DF*. D'autre part, le *NPR* et le *MR* (et non pas le *DF*) signalent les phrasèmes par une série d'indicateurs métalinguistiques antéposés⁶⁸. L'indicateur le plus fréquent et le plus «explicite» est *LOC.*, les autres étant (voir Heinz, 1993) *FIG.*, *PROV.*, *ALLUS.*, *MÉTAPH.*, *MÉTON...* Nous pouvons encore noter, que les phrasèmes se reconnaissent, d'une façon plutôt indirecte, soit par la forme neutralisée sous laquelle ils apparaissent, soit par la présence d'une définition ou d'une glose.

Mais puisque le sens du phrasème est dans une grande mesure idiosyncrasique par rapport aux sens de ses constituants, il est nécessaire, pour qu'il puisse être intégré dans une division de sens, qu'il soit remotivé. Par exemple, pour pouvoir faire figurer *tenir le haut de pavé* dans une division sémantique de l'article de *pavé* (2.), les rédac-

⁶⁶ Mais, peut-être, ce facteur n'est, en effet, qu'une combinaison des autres facteurs évoqués (stabilité d'éléments, longueur d'article)

⁶⁷ Nous nous référons principalement à Heinz, 1993, qui propose une étude exhaustive de la description des phrasèmes dans *Le Petit Robert*, ainsi qu'à Wiegand, 1990

⁶⁸ Le terme *indicateur* est emprunté à M. Heinz, 1993, qui parle plus précisément d'*indicateur de statut locutionnel*.

teurs du *NPR* étaient obligés de remotiver le sens de *pavé* qui, en tant qu'élément du phrasème, n'a plus rien à voir (en synchronie) avec le sens «la partie de la voie publique revêtue de...»⁶⁹. Il peut donc se produire le paradoxe suivant: l'utilisateur qui cherche le sens d'un phrasème devrait, pour se faciliter la recherche, connaître l'étymologie du phrasème et remotiver ses constituants. Il va sans dire qu'un locuteur étranger qui ne connaît pas le sens du phrasème, connaît encore moins son origine. Le placement d'un phrasème à l'intérieur d'un article «intégré» est donc, pour cet utilisateur, totalement (ou partiellement) imprévisible. Prenons comme exemple l'article *pavé* du *NPR* où les phrasèmes sont répartis dans les différentes divisions sémantiques (pour un repérage plus facile nous les avons soulignés de traits onduleux):

1. **PAVÉ** [...] n. m. • 1312; de *paver* 1 ♦ Ensemble des blocs (de pierre, de bois, etc.) qui forment le revêtement du sol. ⇒ pavage, pavement. «à Tolède, où le pavé est composé de petits cailloux polis» (Gautier). *Le pavé de marbre, de mosaïque, d'une église*. 2 ♦ Spécialt La partie d'une voie publique ainsi revêtue. *Pavé humide, glissant*. LOC. Brûler le pavé: courir ou rouler très vite. — (De l'époque où le ruisseau occupait le milieu de la rue) Tenir le haut du pavé: occuper le premier rang. ◇ La rue, la voie publique. LOC. Battre* le pavé. Être sur le pavé, sans domicile, sans emploi. Mettre, jeter qqn sur le pavé (cf. À la rue*). 3 ♦ (XVI^e) Chacun des blocs de basalte, de granit, de grès ou de bois spécialement taillés et préparés pour revêtir un sol. ⇒ aussi carreau, l. dalle. Joints entre les pavés. Poser des pavés. ⇒ paver. Arracher des pavés pour faire une barricade. ◇ LOC. Le pavé de l'ours*. — C'est le pavé dans la mare, un événement inattendu qui apporte la surprise et le trouble dans une situation tranquille. Avoir un pavé sur l'estomac: avoir un poids indigeste. 4 ♦ FIG. Bloc. «Un carré de filet de bœuf, un véritable pavé de viande» (Romains). — Spécialt Épaisse tranche de bœuf. *Un pavé grillé*. *Le pavé de Charolais à la moelle*. ◇ Gâteau en forme de pavé. *Pavé au chocolat*. 5 ♦ FAM. et PÉJ. Livre très épais. *Un énorme pavé de huit cents pages*. — Article de journal imprimé d'une manière massive. *Un pavé publicitaire*. — Texte trop long et lourdement rédigé. *Quel pavé, cette thèse!* 6 ♦ INFORM. Partie séparée (généralement carrée) d'un clavier d'ordinateur comprenant les chiffres et les symboles opératoires. ⊗ HOM. Pavée.

L'intégration des phrasèmes présente des inconvénients sérieux aux utilisateurs en situation de décodage, puisque ces utilisateurs ont besoin d'un accès rapide et aisé à l'information cherchée: cette recherche «rapide» ou «ponctuelle» est d'ailleurs la situation d'utilisation la plus courante. Mais, en compensation, l'intégration se prête mieux à une consultation plus «poussée», à une recherche «globale» (voir Heinz, 1993): l'intégration permet aux utilisateurs de cerner l'étymologie d'un phrasème, de relever les connotations qu'une communauté linguistique associe à certains concepts etc.

⁶⁹ Pour justifier le rattachement, les rédacteurs ont fait précéder le phrasème d'une courte explication étymologique: «De l'époque où le ruisseau occupait le milieu de la rue».

4.5.2.2.2. Annexion. Ce procédé qui regroupe les phrasèmes en une sorte d'annexe placée en fin d'article est rarement pratiqué (au moins, «à l'état pur») en lexicographie française. Il est plus fréquent dans les dictionnaires allemands et anglais. Le *New Oxford Dictionary of English* en est un excellent exemple: les moyens typographiques (passage à la ligne, étiquette «*PHRASES*» au début de l'annexe) séparent nettement l'annexe qui regroupe les phrasèmes du reste de l'article. Parmi les dictionnaires français, ce sont surtout les dictionnaires de la maison Larousse (*Grand dictionnaire encyclopédique Larousse, Lexis*) qui ont recours à une variété de ce procédé. Prenons l'article *pavé* du *Lexis* que nous pouvons comparer avec celui du *NPR* (ci-dessus):

1. PAVÉ [...] n. m. (de *paver*; 1312). 1. Petit bloc de pierre, généralement destiné au revêtement des chaussées ou des routes; [...] – 2. La rue, la voie publique: *Errer sur le pavé de Paris*. – 3. (v. 1500) Éléments d'un carrelage (syn. CARREAU). – 4. *Battre le pavé*, marcher dans les rues, sans but précis. || *Être sur le pavé*, être sans domicile (syn. ÊTRE À LA RUE); être sans emploi: *Elle ne se rendait pas compte qu'elle se ruinait, qu'elle allait se trouver bientôt sur le pavé* (Giono) || *Pavé de l'ours*, acte irréfléchi et funeste inspiré par une amitié aveugle. || *Tenir le haut du pavé*, tenir le premier rang; jouir d'une grande considération. || *Un pavé dans la mare*, une vérité qui jette la perturbation.

Les phrasèmes sont regroupés à la fin de l'article, dans une division à part. Mais il n'y a ni passage à la ligne, ni étiquette (ou symbole) spécifique signalant qu'il ne s'agit pas d'une division sémantique. À l'intérieur de la division, les phrasèmes sont ordonnés alphabétiquement et séparés entre eux par des doubles traits verticaux. Leur statut d'unités de langue est souligné par la forme neutralisée ainsi que par le fait qu'ils constituent des sous-adresses de différentes rubriques (définitions, exemples d'usage).

4.5.3. Un dictionnaire qui s'adresserait aux utilisateurs étrangers devrait, à notre avis, opter impérativement pour l'annexion. La «raison d'être» des phrasèmes dans un dictionnaire monolingue est essentiellement le décodage. Et comme nous l'avons déjà noté, c'est justement l'annexion qui se prête le mieux au décodage puisqu'elle permet une recherche «plus rapide» du phrasème inconnu⁷⁰. Les critères qui décideraient sous quel constituant le phrasème sera traité devraient être formels et cohérents. La solution idéale serait de mentionner le phrasème sous tous les constituants (sauf sous les mots grammaticaux) et de mettre en place un système de renvois qui orienterait l'utilisateur vers l'article où il trouverait une description complète du phrasème cherché (définition, exemple d'emploi, marque d'usage, prononciation...). À l'intérieur de l'annexe nettement détachée du reste de l'article le classement des phrasèmes regroupés devrait obéir à l'ordre alphabétique. On pourrait envisager même deux sortes d'annexes: l'une serait destinée à la description des noms composés et l'autre aux autres types de phrasèmes.

⁷⁰ L'utilisateur qui doit interrompre sa lecture pour consulter un dictionnaire afin de connaître le sens d'un phrasème est un utilisateur pressé, impatient.

L'annexion des phrasèmes, de plus, ne serait pas sans incidences sur le traitement des collocations dont l'intégration dans des divisions sémantiques appropriées, à la différence des phrasèmes, ne pose pas de problèmes. Les collocations et les phrasèmes ne seraient plus décrits aux mêmes endroits de l'article: les premières feraient partie des divisions de sens, les seconds de l'annexe. Cette séparation «formelle» mettrait mieux en évidence les différences, à la fois, de leurs statuts linguistiques et de leurs fonctions lexicographiques.

5. POUR CONCLURE

Les expressions idiomatiques constituent un champ d'analyse linguistique bien vaste et assez hétérogène. Comme nous avons eu l'occasion de le voir, la délimitation du phénomène des expressions idiomatiques ainsi que leur classement requièrent l'intervention de facteurs de nature divergente: sémantique, syntaxique, pragmatique, voire psycholinguistique. Dans notre essai de typologie, nous avons proposé d'opposer aux combinaisons syntagmatiques libres trois types d'expressions idiomatiques: *collocations* (combinaisons récurrentes, imprédictibles à l'encodage) *pragmatèmes* (préconstruits langagiers, figés par rapport à des situations de communication) et *phrasèmes* (expressions polylexicales connaissant un haut degré de figement). L'objectif principal de notre classification, qui ne se veut ni trop rigoureuse ni trop exclusive, n'est pas de dresser une nouvelle, une «nième» classification en tant que telle, mais plutôt de faire ressortir, d'une façon plus systématique, les propriétés linguistiques de différents types d'expressions idiomatiques qui soulèvent une grande diversité des problèmes pour la description lexicographique et auxquels les lexicographes doivent adapter leurs démarches.

Il va sans dire que les problèmes se compliquent davantage dès que l'on ajoute à une analyse «purement» linguistique des considérations lexicographiques, notamment les besoins des différents types d'utilisateurs dans les situations d'utilisation distinctes. Pour notre examen du traitement des trois types d'expressions idiomatiques dans un corpus choisi de dictionnaires monolingues français, nous avons adopté la perspective d'un locuteur/apprenant étranger. Nous avons relevé de nombreuses insuffisances et incohérences qui touchent à plusieurs aspects de leur description lexicographique (sélection, formes d'adressage, placement, dispositifs de présentation) et suggéré des solutions:

- quant à la sélection et à l'arrangement des collocations, un rôle plus important devrait être assigné aux critères d'ordre sémantique; à l'intérieur des divisions sémantiques des articles, distinguer clairement (par des moyens typographiques) les collocations des exemples de combinaisons libres;
- prévoir un texte à part, hors de la macrostructure centrale, où les pragmatèmes seraient ordonnés selon des critères onomasiologiques;
- regrouper les phrasèmes dans une ou plusieurs annexes nettement détachées du reste de l'article; à l'intérieur des annexes, classer les phrasèmes selon des critères sémasiologiques; donner un traitement dictionnaire complet (gloses, exemples d'usage, marques d'usage, prononciation).

BIBLIOGRAPHIE

1. DICTIONNAIRES:

- DF* = *Dictionnaire du français. Référence – Apprentissage* (sous la direction de J. Rey-Debove), Paris, Dictionnaire Le Robert / CLE international, 1999.
- DFC* = *Dictionnaire du français contemporain* (sous la direction de J. Dubois), Paris, Larousse, 1971 (¹1967).
- Lexis* = *Dictionnaire de la langue française: Lexis* (sous la direction de J. Dubois), Paris, Larousse-Bordas, 1999 (¹1979).
- MR* = *Le Micro-Robert. Dictionnaire d'apprentissage de la langue française* (sous la direction de A. Rey), Paris, Dictionnaire Le Robert, 1988 (¹1971).
- NPR* = *Le nouveau Petit Robert* (sous la direction de J. Rey-Debove & A. Rey), Paris, Dictionnaires Le Robert, 1993 (¹1967) [version papier et version sur CD-ROM (Havas interactive, 1997)].
- PL* = *Le Petit Larousse. Compact 1997* (sous la direction de P. Maubourguet), Paris, Larousse, 1996 (¹1905).

2. AUTRES RÉFÉRENCES

- ALONSO RAMOS, M. & Tutin, A. (1996): «A classification and Description of Lexical Functions for the Analysis of their Combinations», in: Wanner éd. 1996, 147-167.
- ASHER, R. E. éd. (1994), *The Encyclopedia of Language and Linguistics*, 10 vol., Oxford, Pergamon Press.
- ATKINS, B. T. S. (1985), «Monolingual and bilingual learners' dictionaries: A Comparison», in: Ilson éd. (1985), 15-24.
- BALLY, C. (1920,¹1909), *Traité de stylistique française*, Vol. 1, Paris, Klincksieck.
- — (1921,¹1909), *Traité de stylistique française*, Vol. 2, Paris, Klincksieck.
- BÉJOINT, H. (2000), *Modern Lexicography: An Introduction*, Oxford, Oxford University Press.
- BÉJOINT, H. & Moulin, A. (1987), «The place of the dictionary in vocabulary acquisition», in: Cowie éd. (1987), 97-114.
- BENSON, M. & Benson, E. & Ilson, R. (1997), *The BBI Dictionary of English Word Combinations*, Amsterdam/Philadelphia, John Benjamins.
- BOGAARDS, P. (1999), «Access structure of learners' dictionaries», in: Herbst & Popp éd. (1999), 113-130.
- COWIE, A. P. (1981), «The Treatment of Collocations and Idioms in Learners' Dictionaries», *Applied Linguistics* 2/3, 223-235.
- — éd. (1987), *The Dictionary and the Language Learner*, Tübingen, Max Niemeyer Verlag.
- — (1994), «Phraseology», in: Asher éd. (1994), 3168-3171.
- — éd. (1998), *Phraseology. Theory, Analysis and Applications*, Oxford, Clarendon Press.
- FIRTH, J. R. (1957), *Papers in Linguistics 1934-1951*, Oxford, Oxford University Press.
- FÓNAGY, I. (1982), *Situation et signification*, Amsterdam, John Benjamins.
- — (1997), «Figement et changement sémantiques», in: Martins-Baltar éd. (1997), 131-164.
- GABROVŠEK, D. (1998), «Coping with *stubborn stains* and *persistent headges* – for What It's Worth: Word Combinability in Action», *Vestnik* 32/1-2, 111-154.
- GROSS, G. (1996), *Les expressions figées en français*, Gap/Paris, Ophrys.
- — (1997), «Du bon usage de la notion de *locution*», in: Martins-Baltar éd. (1997), 201-223.
- GROSS, M. (1981), «Les bases empiriques de la notion de prédicat sémantique», *Langages* 63, 7-52.
- GÜLICH, E. & Kraft, U. (1997), «Le rôle du «préfabriqué» dans le processus de production discursive», in: Martins-Baltar éd. (1997), 241-276.
- HAUSMANN, F. J. (1979), «Le dictionnaire des collocations est-il possible?», *Travaux de linguistique et de littérature* XVII, 187-195.
- — (1989), «Le dictionnaire des collocations», in: Hausmann et al. éd. (1989-1990-1991), 1010-1019.
- — (1990), «Les dictionnaires pour l'enseignement de la langue étrangère: français», in: Hausmann et al. éd. (1989-1990-1991), 1386-1390.
- — (1990a), «De quoi se compose l'article du dictionnaire de langue? L'importance du sous-adresse», in: Magay & Zigány éd. (1990), 59-66.
- — (1999), «Semiotaxis and learners' dictionaries», in: Herbst & Popp éd. (1999), 205-211.
- HAUSMANN, F. J. & Wiegand, H. E. (1989), «Component Parts and Structures of General Monolingual Dictionaries: A Survey», in: Hausmann et al. éd. (1989-1990-1991), 328-360.

- HAUSMANN, F. J. & Reichmann, O. & Wiegand, H. E. & Zgusta, L. éd.s. (1989-1990-1991), *Wörterbücher/ Dictionaries/Dictionnaires. Ein internationales Handbuch zur Lexikographie/An International Encyclopedia of Lexicography/Encyclopédie internationale de lexicographie*, 3 Vol., Berlin/New York, Walter de Gruyter.
- HEINZ, M. (1993), *Les locutions figurées dans le «Petit Robert»*, Tübingen, Max Niemeyer Verlag (Lexicographica Series Maior 49).
- HERBST, T. & Popp, K. éd. (1999), *The Perfect Learners' Dictionary (?)*, Tübingen, Max Niemeyer Verlag (Lexicographica Series Maior 95).
- ILSON, R. F. éd. (1985), *Dictionaries, Lexicography and Language Learning*, Oxford, Pergamon Press (ELT Documents 120).
- KLEIBER, G. (1990), *La sémantique du prototype*, Paris, P.U.F. (Linguistique nouvelle).
- LAMY, M.-N. (1985), «Innovative practices in French monolingual learners' dictionaries as compared with their English counterparts», in: Ilson éd. (1985), 25-34.
- LANGACKER, R. (1987), *The Foundations of Cognitive Grammar*, Vol. 1 (*Theoretical Prerequisites*), Stanford, Stanford University Press.
- MAGAY, Z. & Zigány, J. éd.s. (1990), *BudaLEX '88 Proceedings*, Budapest, Akadémiai Kiadó.
- MANNING, C. & Schütze, H. (1999), *Foundations of Statistical Natural Language Processing*, Cambridge (Mass.), M.I.T. Press.
- MARTIN, R. (1997), «Sur les facteurs du figement lexical», in: Martins-Baltar éd. (1997), 291-305.
- MARTINS-BALTAR, M. éd. (1997), *La locution entre langue et usages*, Fontenay/Saint-Cloud, ENS Éditions.
- MEL'ČUK, I. A. (1993), *Cours de morphologie générale (théorique et descriptive)*, Vol. 1 (*Introduction et 1^{ère} partie: Le Mot*), Montréal/Paris, Les Presses de l'Université de Montréal/CNRS.
- — (1993a), «La phraséologie et son rôle dans l'enseignement/apprentissage d'une langue étrangère», *Études de linguistique appliquée* 92, 82-111.
- — (1996), «Lexical Functions: A Tool for the Description of Lexical Relations in a Lexicon», in: Wanner éd. (1996), 37-102.
- — (1997), *Cours de morphologie générale (théorique et descriptive)*, Vol. 4 (5^e partie: *Signes morphologiques*), Montréal/Paris, Les Presses de l'Université de Montréal/ CNRS.
- — (1997a), *Vers une linguistique SENS-TEXTE. Leçon inaugurale (faite le Vendredi 10 janvier 1997)*, Paris, Collège de France.
- — (1998), «Collocations and Lexical Functions», in: COWIE éd. (1998), 23-53.
- MEL'ČUK ET AL. (1984-1988-1992-1999), *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques. (= DECFC)* Vol. 1-4, Montréal, Les Presses de l'Université de Montréal.
- MEL'ČUK, I. A. & Clas, A. & Polguère, A. (1995), *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve, Duculot.
- Mel'čuk, I. A. & Wanner, L. (1996), «Lexical Functions and Lexical Inheritance for Emotion Lexemes in German», in: Wanner éd. (1996), 209-278.
- MOUNIN, G. (1995,¹1974), *Dictionnaire de la linguistique*, Paris, P.U.F.
- POLGUÈRE, A. (2000), «Une base de données lexicales du français et ses applications possibles en didactique», *LIDIL* 21, 75-97.
- REY-DEBOVE, J. (1971), *Étude linguistique et sémiotique des dictionnaires français contemporains*, The Hague/Paris, Mouton (Approaches to Semiotics 13).
- — (1989), «Le traitement analogique dans le dictionnaire monolingue», in: Hausmann et al. éd.s. (1989-1990-1991), 635-640
- RIEGEL, M. & PELLAT, J.-C. & RIOUL, R. (1998,¹1994), *Grammaire méthodique du français*, Paris, P.U.F.
- SINCLAIR, J. (1991), *Corpus, Concordance, Collocation*, Oxford, Oxford University Press.
- WANNER, L. éd. (1996), *Lexical Functions in Lexicography and Natural Language Processing*, Amsterdam/Philadelphia, John Benjamins.
- WIEGAND, H. E. (1989), «Printed Dictionaries and Their Parts as Texts. An Overview of More Recent Research as an Introduction», *Lexicographica* 6, 1-126.
- Wouden, Ton Van der (1997), *Negative Contexts: Collocation, Polarity and Multiple Negation*, London/New York, Routledge.

APPENDICE

La liste des 125 collocations examinées quant à leur traitement dictionnaire.

Mag: célibataire endurci, intention ferme, prix exorbitant, saleté repoussante, reproches sanglants, activité fiévreuse, blancheur éclatante, amour fou, peur bleue, nuit profonde, profond sommeil, vive douleur, chaleur suffocante, passion dévorante, besoin pressant, injustice criante, terrible injustice, confiance absolue, refus formel, réponse catégorique, regretter vivement, répliquer vivement, surveiller étroitement, dormir profondément, souhaiter/désirer ardemment, infiniment reconnaissant, grièvement blessé, gravement malade, largement ouvert, absolument faux

Mult: bordée d'injures, salve d'applaudissements, tempête d'applaudissements, tonnerre d'applaudissements, torrent(s) de larmes, régime de bananes, banc de poissons, botte de radis, trousseau de clés, meute de chiens, volée d'oiseaux

Sing: accès de colère, accès de fièvre, poussée de fièvre, goutte de pluie

Son: (le) chat miaule, (l') éléphant barrit, (le) cochon grogne, (l') abeille/guêpe bourdonne, (le) cheval hennit, (le) feu crépite, (l') âne brait, (la) poule caquette, (le) renard glapit, (la) porte claque

Oper₁: faire des excuses, détenir le pouvoir, détenir le record, faire une remarque, faire (une) erreur, commettre une erreur, commettre une imprudence, commettre un crime, perpétrer un crime, commettre un méfait, donner un ordre, donner (un) conseil, faire une comparaison, déposer une plainte, faire un procès, dresser une liste, établir un impôt, établir des liens, lancer un défi, lancer un ultimatum

Oper₂: courir un danger, courir un risque, subir une opération, passer un examen, subir une défaite, se heurter à un refus, figurer sur la liste

Func₀: (le) silence règne, (le) vent souffle

Func₁: la difficulté réside, (le) malheur frappe

Real₁: résoudre un problème, tenir la promesse, assouvir la curiosité, tendre un piège, succomber à la tentation, infliger une amende

Real₂: purger une peine, réussir à un examen, être reçu à un examen, suivre un conseil, tomber dans un piège

Real₃: exécuter un ordre

Fact₀: (le) souhait s'accomplit

Incep(...): prendre forme, tomber malade, tomber amoureux, s'armer de patience, (le) sommeil s'empare [de N], (le) sommeil envahit [N], intenter un procès, engager une conversation, entamer des négociations, engager des négociations,

Fin(...): (l') espoir s'envole, dissiper un malentendu, (le) souvenir s'éteint, (la) pensée quitte [N],

Caus(...): attirer l'attention [sur N], exciter l'imagination, susciter l'intérêt, soulever l'enthousiasme, poser un problème, créer des ennuis

Liqu: dissoudre l'assemblée, surmonter la peur,

(...)**Manif:** la vérité jaillit, (la) colère explose, formuler un souhait, présenter des excuses

IDIOMI: TEORETIČNI OPIS IN SLOVARSKA OBRAVNAVA
(PRIMER FRANCOSKIH ENOJEZIČNIH SLOVARJEV)

Članek je v svojem prvem delu osredotočen na jezikoslovno-teoretično opredelitev leksikaliziranih besednih zvez, idiomov (fr. *expressions idiomatiques*). Pregledno so predstavljeni različni kriteriji, ki jih jezikoslovje običajno pritegne v razmejitve med idiomom in svobodno grajenimi besednimi zvezami (pomenski, skladenjski, pragmatični, psiholingvistični kriteriji), ter orisane težave, ki izvirajo iz njihove heterogenosti in stopenjskosti. Idiomi so v nadaljevanju členjeni na sopojavljanke (fr. *collocations*), pragmateme (fr. *pragmatèmes*) in frazeme (fr. *phrasèmes*). Pri predstavitvi tipologije je bilo posebej poudarjeno, da je na posamezne vrste idiomov primerneje gledati kot na »prototipske« kategorije brez (pre)ostrih razmejitev. Vsekakor pa je tipologija koristna, saj nudi sistematičen vpogled v ključne jezikoslovne značilnosti različnih vrst idiomov, ki jih morajo pri slovarskem opisu upoštevati sestavljavci slovarja.

Drugi del ponuja kritičen pretres slovarske obravnave idiomov v izbranem korpusu enojezičnih slovarjev. Korpus je sestavljen skladno z odločitvijo, da bo analiza upoštevala izključno specifične potrebe tujejezičnega govorca – tako so bili izbrani le enozvezkovni pedagoški in splošni slovarji, ki so namenjeni tudi ali pa predvsem tujejezičnim uporabnikom (NPR, MR, Lexis oziroma DF). Analiza se je v več ločenih podglavjih osredotočila na različne aspekte slovarskega opisa posameznih vrst idiomov: na kriterije njihove izbire, na mesto obravnave v slovarskem besedilu in v besedilu slovarskega članka, na razporeditev, na obliko predstavitve ter na vrsto in bogastvo informacij, ki jih mora podati slovar. Članek je opozoril na nekatere težave in pomanjkljivosti, predlagane pa so tudi rešitve, ki se zdijo ustrezne za slovar, ki naj bi ga uporabljali tujejezični govorniki.