

- Boyd-Graber, Jordan. 2002. Collocation Tagging. *Caltech Summer Undergraduate Fellowship — Final Report*. <http://jbg.caltech.edu/surf02/final.pdf>.
- Hanneforth, Thomas. 2002. *Eigennamenerkennung mit lokalen {G}rammatiken*. Vortrag Universität Heidelberg.
- Karttunen, Lauri, J-P. Chanod, G.Grefenstette, A. Schiller. 1996. *Regular Expressions for Language Engineering*. *Natural Language Engineering* 2 (4) 305–328.
- Klavans, Judith. 1996. *Computing Corpus and Machine-Readable Dictionary Data for Building Bilingual Lexicons*. *Machine Translation* 10: 3–4, 1–34.
- Lezius, Wolfgang; Rapp, Reinhard and Manfred Wettler. 1996. *A Morphology-System and Part-of-Speech Tagger for German*. In D. Gibbon, ed., *Natural Language Processing and Speech Technology*. Results of the 3rd KONVENS Conference, pp. 369–378. Mouton de Gruyter.
- Schemann, Hans. 1993. *Deutsche Idiomatik. Die deutschen Redewendungen im Kontext*. Stuttgart: Pons.
- Schiller, Anne, Teufel, Simone, Stöckert, Christiane and Christine Thielen. *Vorläufige Guidelines für das Taggen deutscher Textcorpora mit STTS*. Technical report, IMS, Univ. Stuttgart and Sfs, Univ. Tübingen, 1995.
- Senellart, Jean. 1998. *Reconnaissance automatique des entrées du lexique-grammaire des phrases figées*. In Lamiroy, B. (éd.) *Le lexique-grammaire*. *Travaux de Linguistique*, 37, 109–127.
- Tschichold, Cornelia. 2000. *Multi-word units in natural language processing*. Zürich: Olms.

### Summary

This work demonstrates the assignment of multi-word expressions in print dictionaries to POS classes with minimal linguistic resources. In this application, 32,000 entries from the *Wörterbuch der deutschen Idiomatik* (H. Schemann 1993) were classified using an inductive description of POS sequences in conjunction with a Brill Tagger trained on manually tagged idiomatic entries. This process assigned categories to 86% of entries with 88% accuracy. This classification supplies a meaningful preprocessing step for further applications: the resulting POS-sequences for all idiomatic entries might be used for the automatic recognition of multi-word lexemes in unrestricted text.

### Authors' addresses

Alexander Geyken  
Berlin-Brandenburgische Akademie der  
Wissenschaften  
Project DWDS  
Jaegerstr. 22/23  
D-10117 Berlin  
geyken@bbaw.de

Jordan Boyd-Graber  
MSC 388  
Caltech  
Pasadena, CA 91126  
jbg@caltech.edu

Reçu le 17 novembre 2003

## Figement et variations en français de Belgique, de France, du Québec et de Suisse\*

Jean René Klein et Corinne Rossari  
Université de Louvain / Université de Fribourg<sup>1</sup>

### 1. Introduction

Dans cet article, nous évoquerons différents problèmes relatifs à la notion de figement, qui émergent dès que l'on se trouve face à la comparaison entre diverses variétés de français. Ainsi, nous aborderons la notion de figement d'un point de vue empirique en faisant ressortir la pluralité des facettes qu'elle recouvre, dès qu'elle est mise en relation avec la notion de variation. Notre contribution est, en effet, issue de réflexions menées dans le cadre du projet *BFQS*,<sup>2</sup> qui consiste en la rédaction d'un dictionnaire des locutions figées de quatre variétés de français (Belgique, France, Québec, Suisse). La masse considérable de données que nous avons à traiter dans le cadre de ce projet induit la particularité de l'approche que nous allons adopter face à la notion de figement.

### 2. Le figement : rappel de quelques traits

Avant d'entamer l'analyse de cas concrets, nous rappelons brièvement des traits inhérents à la notion de figement que nous utiliserons pour notre description. Outre les trois traits constitutifs des expressions figées (EF) par rapport aux expressions libres (EL), dégagés par B. Lamiroy (2003:7),

	EF	EL
opacité sémantique	+	–
paradigme limité	+	–
restrictions morphosynt.	+	–

deux phénomènes sont essentiels à la compréhension du phénomène. En premier lieu, on notera le caractère gradué du figement, souligné par de nombreux auteurs (M.Gross 1988, S.Mejri 1997, Ch.Schapira 1999, F.Tollis 2001), et très nettement, par G.Gross (1997:213) et D.Gaetone (1997:169). L'exemple du verbe *prendre* et de ses compléments possibles en est une parfaite illustration. *Prendre*, dans son sens prototypique de 'saisir qqch ou qqn, généralement avec une partie du corps ou avec un instrument', admet des objets innombrables, ce qui ne signifie pas indifférents. Dès que l'on s'écarte de ce sémantisme basique, le nombre d'objets possibles décroît aussitôt et l'on constate que le verbe régit une classe d'objets, ce qui a pour effet d'engendrer des séquences plus figées.

- (1) a. *Prendre (un train)*  
(emprunter un moyen de communication)  
b. *Prendre (un verre, un repas)*  
(absorber, ingérer)  
c. *Prendre (un chemin, une rue)*  
(suivre, s'engager dans)

En second lieu, intervient ce que R.Martin (1997:295) appelle le phénomène de sélectivité normative, c'est-à-dire que la norme<sup>3</sup> opère un choix parmi les diverses possibilités théoriquement offertes par la langue. Par exemple, à l'intérieur du français commun<sup>4</sup> les restrictions suivantes s'opèrent :

- (2) (BFQS) *Cela a pesé (lourd + \*léger) dans sa décision*  
(BFQS) *Max en sait (long + \*court) sur le sujet*

De même, en comparant le français commun avec une variété, on remarque que la sélection ne se fait pas de la même manière :

- (3) (B) *On mangera un bout ensemble à midi*  
(FQS) *On mangera un \*bout ensemble à midi*  
(BFQS) *On mangera un morceau ensemble à midi*  
(BFQS) *Max en sait un (bout + \*morceau) sur le sujet*

Ceci relève bien de l'arbitraire de la norme, car le choix existe en français commun, par exemple, dans les séquences libres telles un *morceau* (un *bout*) de pain.

La comparaison entre différentes formes de figement, selon les variétés, attire donc l'attention sur des faits de sélection normative.

### 3. La perspective comparative et les cas de figement

Comment la comparaison entre variétés du français fait-elle ressortir ces cas? Nous proposons ci-dessous d'examiner la valeur heuristique de la comparaison de séquences où apparaissent des traces de figement

- (4) (S) *Pour ouvrir la porte, il faut peser sur ce bouton*

Nul n'aurait tendance à considérer un syntagme comme *appuyer sur un bouton* comme une expression figée (EF). Il n'y a en effet aucun des trois traits indiqués plus haut, comme constitutifs du figement :

- *opacité sémantique*: tous les termes du groupe possèdent leur sens (dans d'autres groupes *appuyer* et *bouton* peuvent avoir le même type de dénotation);
- *paradigme limité*: *bouton* n'est pas un exemplaire d'un paradigme limité: le verbe *appuyer* peut avoir comme objet toute une série de N [-hum].
- *restriction morphosyntaxique*: il n'y en a aucune: on peut modifier le déterminant, conjuguer le verbe à tous les temps... *Appuie sur ce bouton!*

Dans la variété suisse, le verbe *peser* correspond à *appuyer* dans certains emplois. On peut donc utiliser la séquence *peser sur un bouton*. En comparant les combinaisons possibles avec *peser* et *appuyer*, on s'aperçoit que *peser sur un bouton* est plus restreint :

- on peut appuyer avec n'importe quoi: ses pieds, ses mains, ses doigts, un bâton...
- on peut appuyer sur n'importe quoi: sur le plancher, sur le plafond, sur un stylo-feutre qui n'a plus beaucoup d'encre...

Le groupe *peser sur un bouton* qui est bien l'équivalent sémantique de *appuyer sur un bouton* n'a pas toutes ces possibilités :

- on pèse forcément avec ses doigts: \*Pèse avec le bâton sur le bouton; \*Pèse avec tes pieds sur la porte
- on ne pèse pas sur n'importe quoi: \*Pèse sur le stylo-feutre s'il n'a plus d'encre; \*Pèse fort sur le mur si tu veux que ça colle
- Si l'on pèse sur autre chose que sur des boutons: il s'agit alors de cas d'ellipse. (S) Pèse sur rez pour arriver au demi sous-sol; (S) Pèse sur la sonnette un coup avant d'entrer.

C'est donc le facteur *paradigme limité* qui est responsable du caractère figé du groupe et qui ressort par la comparaison entre l'expression du français standard<sup>5</sup>

et l'expression (S). En (S), le verbe *peser* peut signifier 'appuyer avec ses doigts sur un bouton à fonction spécifique (ascenseur, sonnette)'. Cet emploi spécialisé est soumis à des variations régionales. Dans certaines régions (le canton de Vaud), *peser* a un emploi moins restreint et tend à devenir l'équivalent de *appuyer*. Si cette tendance se généralisait en français de Suisse, il n'y aurait plus lieu de considérer le groupe comme une expression figée. La comparaison entre la classe d'objets qui admet *appuyer* en français et celle permise par *peser* dans la variété suisse met donc en lumière une contrainte manifestant un cas de figement.

Un cas symétrique est l'emploi des verbes *éteindre* et *allumer* qui, en (S), est beaucoup plus étendu qu'en français commun.

(5) (S) *Luc a oublié d'éteindre le robinet*

En français standard, on ne peut allumer ou éteindre que des objets incandescents (une lumière, un feu, des phares, une torche et par analogie le radiateur, le moteur, la télévision), mis à part les emplois métaphoriques (allumer une fille, un garçon). En (S), l'emploi de ces deux verbes s'est étendu à certaines acceptions des verbes *ouvrir* et *fermer*. On peut allumer ou éteindre l'eau, un robinet. On a affaire à une extension de la classe d'objets en (S), qui peut être sentie comme abusive, selon la norme du français standard, ou comme une forme d'associations préférentielles, dans une perspective comparative; on met ainsi le doigt sur une combinaison dont la sémantique s'écarte de la classe d'objets possibles en français standard.

Sans la perspective comparative, on n'aurait donc pas de raison de repérer *peser sur un bouton* ou *allumer un robinet* comme des expressions figées. *Peser sur un bouton* serait analysé comme illustrant un des emplois du verbe *peser* qui fait partie du lexique suisse, dont l'équivalent en français standard est *appuyer*, et *allumer un robinet*, comme une extension abusive du verbe *allumer*, employée surtout en Suisse.

La variété belge illustre un cas comparable d'association préférentielle. Le français standard connaît *crever* et *claquer* au sens de 'mourir', d'où les emplois connus:

- (6) a. (BFQS) *Max crève de (faim + soif + froid + chaud)*  
 b. (BFQS) *Max crève de (jalousie + rage + rire).*  
 c. (B) *Max claque de (faim + soif + froid + chaud)*  
 d. (B) *Max claque de (\*jalousie + \*rage + \*rire)*

En français de Belgique, *claquer* peut se substituer à *crever*, mais pour une part seulement du paradigme, les compléments renvoyant aux sentiments étant exclus avec *claquer*.

Avec le dernier cas illustré sous (7), la comparaison entre variétés montre des variations affectant simultanément plusieurs constituants d'une expression. *Dix-heures* 'collation prise au milieu de la matinée' est attesté en BS, alors que *quatre-heures* 'collation prise au milieu de l'après-midi' est connu en BFS. Jusque-là, il ne s'agit que d'une variation strictement lexicale. Toutefois, l'attestation de séquences comme:

- (7) a. (BS) *Max prend (un + son) dix-heures*  
 b. (BFS) *Max prend (un + son) quatre-heures*  
 c. (S) *Max (prend + fait) (un + des) dix-heures*  
 d. (S) *Max (prend + fait) (un + des) quatre-heures*  
 e. (S) *Max prend de bonnes dix-heures*

illustre quatre types de variations (sur le prédicat, le complément, le nombre et le genre):

1. *prendre* (un repas, une collation) est le verbe attendu en français commun, alors que pour (S), le verbe *faire* convient aussi (7c-d);
2. la variété (F) ne connaît que *quatre-heures*, alors que (BS) emploient *dix-heures* et *quatre-heures* (7a-b);
3. pour (BF) le déterminant est toujours du singulier, en (S) le pluriel semble plus usuel (7c);
4. enfin, en (S), *dix-heures*, au pluriel, peut être employé au féminin (7e).

Ainsi, ce qui n'apparaissait que peu figé, au départ, en français commun (prendre un certain type de collation) manifeste des contraintes dans la variété (S), qui mettent mieux en évidence une possibilité de figement. Autrement dit, l'étude des variations acquiert une valeur heuristique.

#### 4. Les différentes facettes du figement

##### 4.1 La facette sémantique

C'est une des facettes les plus prototypiques du figement. Une expression formée avec les mêmes mots reçoit des sens différents dans certaines variétés de français. C'est très courant avec les noms de consommations.

- (8) *Luc a bu un demi*  
 (BFQ: Luc a bu 25 ou 33 cl de bière)  
 (S: Luc a bu 50 cl de vin)

Ainsi, *boire un demi* signifie en (BFQ) boire 25cl ou 33cl de bière et en (S) boire un demi-litre de vin. On constate que *demi* ne peut se combiner avec n'importe quel verbe (ainsi, on ne trouve pas *acheter* ou *vendre un demi*, alors qu'on rencontrera *acheter* ou *vendre une bière ou du vin*). Si la différence entre *prendre un demi* en (S) et *prendre un demi* en (BFQ) se limitait à une différence de dénotation, elle pourrait être assimilée à un cas de faux amis. Or, le paradigme avec lequel ce nom se décline n'est pas le même dans ces variétés. Quand *demi* désigne une quantité de bière, il peut se combiner avec un certain nombre de verbes qui servent aussi de prédicats au nom *bière*, alors que quand il réfère, en (S), à un demi-litre de vin, il ne se décline qu'avec des verbes relatifs à la consommation. Ainsi, on ne trouvera en (S) ni «accepter un demi», ni «avalé un demi», alors que ces séquences seront admises avec *demi* signifiant 'bière' en (BFQ).

Dans le même type de contexte, on notera en français de Suisse des emplois elliptiques, strictement contextualisés, de *deux* et *trois* (*décilitres de vin*) constituant des expressions figées totalement opaques pour les usagers des autres variétés:

- (9) a. (S) *Max (prend + boit + commande) deux*  
 b. (S) *Max (prend + boit + commande) trois*

Ces cas ne doivent à nouveau pas être confondus avec ce qui pourrait sembler similaire à première vue, soit une classe d'objets 'boissons' dans des énoncés tels ceux de (10) où l'on trouve des régionalismes lexicaux susceptibles d'apparaître dans d'autres constructions syntaxiques, comme *acheter du péquet*.

- (10) a. (B) *Max prend un péquet*  
 b. (F) *Max prend une verte*  
 c. (Q) *Max prend une liqueur (= boisson gazeuse)*  
 d. (S) *Max prend une petite bleue*

Un autre cas tout à fait exemplaire est celui des expressions construites avec le mot *famille*. Les énoncés de la série (11) permettent l'interprétation analytique à partir du sens standard de *famille* 'les parents, les personnes apparentées':

- (11) a. (BFQS) *Max attend de la famille*  
 b. (BFQS) *Max a de la famille*  
 c. (BFS) *Max est en famille*

Toutefois, pour les mêmes expressions, les diverses variétés connaissent des emplois non compositionnels:

- (12) a. (S) *Léa attend de la famille*  
 (Léa est enceinte)  
 b. (S) *Léa a de la famille*  
 (Léa a des enfants)  
 c. (Q) *Léa est en famille*  
 (Léa est enceinte)

On observe d'autre part des cas déviants en français standard:

- (13) a. (Q) *Léa part pour la famille*  
 (Léa tombe enceinte)  
 b. (B) *Léa attend famille*  
 (Léa est enceinte)  
 c. (B) *Max est famille avec Luc*  
 (Max est parent avec Luc)

#### 4.2 La facette morpho-syntaxique

Il y a aussi des variations de sens plus complexes, car portant sur des mots grammaticaux, tels la préposition.

On trouve en (F), mais cela semble se répandre aussi en (B), la collocation *aller sur N ville* comme *aller sur Paris, sur Rome*, avec le sens de *aller à*. En (S), l'expression n'est pas utilisée pour les villes mais pour certains pays et elle n'a pas exactement la même dénotation: *aller sur N pays* peut signifier 'aller en direction de l'Italie', mais également 'se rendre en Italie', ou encore 'faire un séjour en Italie'.

- (14) a. (S) *Luc va sur l'Italie*  
 (Luc va faire un séjour en Italie)  
 b. (S) *Luc va sur France ce week-end*  
 (Luc va passer le week-end en France)  
 c. (BF) *Luc va sur Paris*

Cette variation concerne le paradigme avec lequel se combine la collocation *aller sur* dont la dénotation n'est pas équivalente en (S) et en (F). Avec *France* (cf. 14b), l'expression a le sens plus large 'aller en France' (pour faire ses courses, une balade, passer des vacances, un week-end) et on remarquera la contrainte que constitue l'absence de déterminant.

Ces cas présentent les particularités suivantes :

- certaines collocations sont liées à des facteurs exclusivement pragmatiques et non explicables (il s'agit des pays où les Suisses se rendent le plus souvent, mais qui ne sont pas forcément limitrophes) ; on entend *aller sur l'Espagne pour les vacances d'été*.
- les constructions ne se déclinent pas toutes selon la même structure. À côté d'*aller sur l'Italie*, on dit *aller sur France* mais \**aller sur la France* est agrammatical pour dire *aller en France*. Certains traits de figement apparaissent ainsi aux niveaux de l'opacité sémantique (la dénotation de *aller sur*), des restrictions morpho-syntaxiques (le déterminant zéro ou défini) et du paradigme limité (seuls quelques pays et villes permettent ce type de collocation).

### 4.3 La facette lexicale

4.3.1 Le français commun a l'expression *avoir du retard* qui est une variante de *être en retard* ou *retarder*, selon le sujet.

(15) (BFQS) (*Paul + le train + la montre*) a du retard

(16) (QS) *La montre a de l'avance*

En (15), le sujet peut être un N ± hum : *Paul a du retard, il ne sait toujours pas lire; Paul a du retard, il est déjà huit heures; ce train a du retard (= ce train est en retard); ma montre a du retard (= ma montre retarde)*. L'expression *avoir du retard* a donné lieu à une extension avec *avance*, qui ne s'est pas stabilisée de la même manière dans les quatre variétés de français. Comme le montre (16), en (Q) et en (S), *la montre a de l'avance* est un équivalent du français commun *la montre avance*. Les variétés (Q) et (S) ont donc étendu les emplois de *avoir du retard* à *avoir de l'avance*. En (B) et en (F), un tel emploi de *avoir de l'avance* semble moins fréquent qu'en (QS) : ainsi, l'emploi de *la montre a de l'avance* pour *la montre avance* n'est pas entré dans l'usage de ces variétés. On voit donc des degrés de stabilisation varier en fonction des variétés.

Une autre forme d'extension intéressante pour la notion de figement est l'utilisation du sens d'un mot appartenant à une certaine catégorie lexicale pour un mot qui a la même base morphologique, mais qui appartient à une autre catégorie lexicale. Le cas pris en exemple concerne de nouveau le verbe *avancer* et le substantif *avance*.

(17) a. (BFQS) *Cela n'avance à rien*

b. (B) *Il n'y a pas d'avance*

En français commun, le verbe *avancer* a le sens de 'mener, conduire' dans l'expression *n'avancer à rien*. La variété (B) a l'expression *il n'y a pas d'avance* qui équivaut à *cela n'avance à rien*. Dans cette expression, le substantif emprunte son sens à celui du verbe *avancer* dans l'expression parallèle. En effet, les deux structures sont similaires : la polarité négative<sup>6</sup> (on ne trouve ni *cela avance à qqc* ni *il y a de l'avance*) et la complétive facultative (*cela n'avance à rien de faire du bruit*, (B) *il n'y a pas d'avance à faire du bruit*). Le parallélisme structurel accentue l'effet de correspondance sémantique entre le verbe et le substantif.

4.3.2 Dans certains cas, la variation lexicale affecte le modifieur. On trouve l'expression *compte dessus et bois de l'eau* associée librement ou obligatoirement à différents modifieurs selon les variétés.

- (18) a. (S) *Compte dessus et bois de l'eau chaude*  
 b. (B) *Compte dessus et bois de l'eau (E + claire)*  
 c. (F) *Compte dessus et bois de l'eau (E + fraîche)*

En (S), l'expression est obligatoirement associée à *chaude*. En (F), elle est facultativement associée au modifieur *fraîche*. En (B), elle est facultativement associée au modifieur *claire*. Toutes ces variations ne modifient pas le sens des différentes expressions qui signifient toutes 'ne pas devoir compter sur qqc'. Ces variations mettent en avant le caractère figé de l'expression de par l'opacité sémantique qu'elles rendent manifeste. Les adjectifs varient sur des paradigmes différents, température située à des pôles opposés de l'échelle (*fraîche, chaude*), qualité (*claire*), alors que le sens ne change pas.

4.3.3 Toute variété de français possède des statalismes, à savoir des expressions issues de ses institutions économiques, politiques, sociales.

- (19) a. (S) *Paul a une casco complète*  
 b. (B) *Paul a une omnium*  
 c. (BFQS) *Paul a une assurance tous risques*

Ainsi, quand on possède pour sa voiture une assurance tous risques, on dit que l'on a une casco complète en (S), une omnium (en B), et une assurance tous risques en (F), qui est l'expression du français commun. Malgré le caractère régional de ces termes administratifs (*casco, omnium*), qui les rend opaques pour les usagers des autres variétés, nous considérons que les structures dans lesquelles ils s'intègrent ne sont pas figées, dans la mesure où il s'agit de constructions diverses et libres, telles que *prendre une casco complète (une*

*omnium*), être assuré avec une casco complète (une *omnium*), conduire avec une casco complète (une *omnium*), etc.

Il existe par contre des sigles qui prennent de l'autonomie. Cela veut dire qu'ils sont intégrés à des expressions où on peut leur associer une dénotation, certes liée à la fonction institutionnelle qu'ils représentent, mais à laquelle s'ajoute une certaine valeur métaphorique. C'est le cas du sigle *A. I.* qui désigne les rentes payées par l'assurance invalidité en Suisse, pour les personnes ayant le statut juridique de handicapés.

- (20) a. (S) *Luc est à l'A. I. depuis son accident*  
 b. (S) *Luc est bon pour l'A. I.*

Ce sigle est interprété en (20a) de manière littérale comme une sorte de contraction de *être à charge des rentes de l'assurance invalidité*. Dans ce cas, l'opacité sémantique se dissipe sitôt que le référent du sigle est décodé. Mais on peut aussi dire *Luc est bon pour l'A. I.* si l'on veut dire de Luc qu'il est au bout du rouleau. Dans ce cas, l'accès au référent du sigle semble presque secondaire pour comprendre le sens de la structure. On a donc un glissement métaphorique, responsable de l'opacité sémantique de la structure, qui lui donne un caractère figé. On le remarque plus clairement en contrastant l'expression avec d'autres qui seraient formées également avec des sigles. Par exemple, le sigle *A. V. S.* désigne les rentes perçues par la caisse des retraites. On peut dire *Paul est à l'A. V. S.* pour dire qu'il est bénéficiaire d'une rente de retraite, comme on peut dire *Paul est à la retraite*, mais on ne peut pas dire *Paul est bon pour l'A. V. S.*, alors que l'on peut dire que *Paul est bon pour la retraite*, si l'on veut signifier que *Paul n'est plus bon à rien dans son travail*. Les deux sigles sont considérés comme faisant partie du lexique suisse (A. Thibault et P. Knecht 1997), mais *A. I.* est seul à avoir acquis une valeur métaphorique dans une expression déterminée.

## 5. Conclusion

Nous avons passé en revue différentes manifestations de ce qui fait ressortir le caractère figé d'une expression en nous basant, entre autres critères, sur la manière dont la norme a stabilisé l'usage dans différentes variétés du français. Cela nous a permis de voir que l'impression de figement peut tenir à toutes sortes de phénomènes, de la variation de genre ou de nombre (*prendre un (des) bon(nes) dix-heures*), à la variation qui touche à la dénotation même d'une expression (*avoir de la famille, être en famille*) ou encore à sa formation morpho-syntaxique,

avec des constructions (non attestées), déviantes selon la norme du français standard (*attendre famille*).

## Notes

\* Nous remercions Béatrice Lamiroy avec qui nous avons eu de nombreuses conversations amicales dont bénéficie la présente contribution.

1. En traitant de la question du figement qui concerne également des formations adverbiales dont la fonction est celle de connecteurs dans le discours, cet article apporte une contribution au projet financé par le Fonds National Suisse No 610-062821.00 dont le but est la constitution d'une typologie sémantique et d'une classification des emplois des connecteurs du français.

2. Pour une présentation du projet, voir: Lamiroy, Leclère, Klein, Labelle (2003).

3. La notion de *norme* est prise ici dans son sens descriptif et non prescriptif.

4. Le terme français commun est utilisé pour désigner le français partagé par les quatre communautés.

5. Nous utilisons la notion de français standard de façon différente de celle de français commun. Le français standard est un français ressenti comme correspondant à la norme du français écrit appris à l'école. Il s'agit d'une notion théorique, alors que le français commun, rappelons-le, représente la langue partagée par les quatre communautés envisagées, indépendamment de toute notion de norme.

6. On notera qu'en français commun la forme interrogative se rencontre aussi: *à quoi cela t'avance-t-il?*

## Références

- Bally, Charles. 1921. *Traité de stylistique française*. Paris: Klincksieck.
- Gaotone, David. 1997. La locution: analyse interne et analyse globale. In: M. Martins-Baltar, (éd.) *La locution entre langue et usages*. Fontenay/Saint-Cloud: ENS éditions, 165-177.
- Gross, Gaston. 1996. *Les expressions figées en français. Noms composés et autres locutions*. Paris/Gap: Ophrys.
- Gaston, Gaston. 1997. Du bon usage de la notion de locution. In: M. Martins-Baltar, (éd.) *La locution entre langue et usages*. Fontenay/Saint-Cloud: ENS éditions, 201-225.
- Gross, Maurice. 1982. Une classification des phrases "figées" du français. *Revue Québécoise de Linguistique*, 11.2, 151-185.
- Gross, Maurice. 1988. Les limites de la phrase figée. *Langages*, 90, 7-22.
- Klein, Jean René; Béatrice Lamiroy 1995. Les expressions figées du français de Belgique. In: J. Labelle, & C. Leclère (éds). *Lexiques-grammaires comparés en français*, Amsterdam/Philadelphia: J. Benjamins, 37-52.

- Kübler, Nathalie. 1995. Parle voir suisse! Présentation d'expressions figées en suisse. In: J. J. Labelle, (éd.). *Lexiques-grammaires comparés et traitements automatiques*. Montréal: Presses de l'UQAM, 267–283.
- Labelle, Jacques. 1988. Lexiques-grammaires comparés: formes figées en français du Québec. *Langages*, 90.
- Lamiroy, Béatrice. 2003. Les notions linguistiques de figement et de contrainte. *Linguisticae Investigationes* XXVI: 1, Amsterdam/Philadelphia: John Benjamins, 1–14.
- Lamiroy, Béatrice; Christian Leclère; Jean René Klein; Jacques Labelle. 2003. Expressions verbales figées et variation en français: le projet BFQS. *Cahiers de lexicologie*, 83.2, 153–172.
- Martin, Robert. 1997. Sur les facteurs du figement lexical. In: M. Martins-Baltar, (éd.) *La locution entre langue et usages*. Paris: ENS éditions. Fontenay-St Cloud, 291–305.
- Martins-Baltar, Michel. (éd.). 1997. *La locution entre langue et usages*. Paris: ENS éditions. Fontenay-St Cloud, 3 vol.
- Mejri, Salah. 1997. *Le figement lexical. Descriptions linguistiques et structuration sémantique*. Tunis: Publications de la faculté des lettres de la Manouba.
- Schapiro, Charlotte. 1999. *Les stéréotypes en français. Proverbes et autres formules*. Paris: Ophrys.
- Thibault, André; Pierre Knecht, (dir.). 1997. *Dictionnaire suisse romand. Particularités lexicales du français contemporain*. Genève: Zoé.
- Tollis, Francis. (éd.). 2001. *La locution et la périphrase. Du lexique à la grammaire*. Paris: L'Harmattan.

### Summary

Set phrases and variation in the French of Belgium, France, Quebec and Switzerland

This article analyses variation in some set phrases in four French varieties: the French of Belgium, France, Quebec and Switzerland. The comparative perspective adopted reveals many types of constraints on the idiomatic properties of the set phrases. The article also demonstrates how observation of this variation can act as a heuristic for the study of theoretical problems related to the definition of the notion of "set phrase" itself.

### Adresses des auteurs

Jean René Klein  
Département d'études romanes  
Centre d'étude sur les lexiques romans  
(CELEXROM)  
Place Blaise Pascal, 1  
B-1348 Louvain-la-Neuve  
klein@rom.ucl.ac.be

Corinne Rossari  
Département de français  
13 av. Beaugard  
1700 Fribourg  
Corinne.Rossari@lettres.unige.ch

## Vers la constitution du lexique-grammaire des verbes en grec moderne

Tita Kyriacopoulou et Anastasia Yannacopoulou  
Université Aristote de Thessaloniki / Université de Marne-la-Vallée

### Introduction

Notre intention étant d'aboutir à une description exhaustive du grec en vue d'une analyse automatique de la langue, la description morphologique des mots doit être complétée par une description syntaxique, voire sémantique. Pour la description syntaxico-sémantique des mots (verbes, noms, etc.) nous avons adopté le modèle du lexique-grammaire élaboré au LADL.<sup>1</sup> Le présent article se limite à la classification des verbes grecs.

Le lexique-grammaire est divisé en tables. Chaque table est caractérisée par une structure type qui définit l'emploi de ses éléments (dans notre cas des verbes). En pratique, il faut étudier les constructions de chaque verbe du dictionnaire morphologique et ensuite réunir dans une table les verbes qui ont une même construction caractéristique. La structure commune de cette construction des verbes définit la table qui les regroupe.

Ce travail, qui constitue la base de la description syntaxique des verbes, est fait par une équipe de trois personnes (T. Kyriacopoulou, R. Voskaki et A. Yannacopoulou). Cette opération de longue haleine a débuté il y a deux ans et n'est pas encore terminée. Nos résultats sont donc incomplets et limités aux verbes pour lesquels les propriétés de structure ont été étudiées.

L'objectif de cet article est de présenter l'organisation générale du lexique-grammaire des verbes et les propriétés qui jouent un rôle crucial dans la répartition des verbes dans les différentes tables de ce lexique-grammaire. Après un bref rappel des principes généraux pris en compte dans notre recherche, nous parlerons d'une première classification des verbes grecs ainsi que des propriétés syntaxiques considérées. Enfin, nous tenterons l'analyse automatique d'un texte en partant d'une table du lexique-grammaire.