

COLLOCATION – NOTION À EMBARRAS TERMINOLOGIQUE

Jana Brňáková
Université d'Ostrava

jana.brnakova@osu.cz

Resumé. La « *collocation* » constitue l'une des notions clé de la linguistique de corpus ainsi que de la phraséologie langagière. Définie en général à l'aide des critères du figement, de fréquence d'utilisation et de la cooccurrence lexicale, sa discrimination dans le cadre du contexte reste douteuse ainsi que l'interprétation de son rôle dans le cadre de la théorie linguistique. Le seul point commun sur lequel s'accorde la majeure partie des spécialistes du domaine est le fait que la mise en place d'une collocation est régie par les propriétés lexico-grammaticales inhérentes et propres à chaque langue naturelle à part.

Mots clés. Collocation. Colligation. Attraction mutuelle. Figement lexical. Figement formel. Fréquence d'utilisation. Voisinage. Cooccurrence lexicale. Lexico-grammaire. Polylexicalité. Phraséologie.

Abstract. Collocation – Terminologically Ambiguous Notion. Collocation, phraseology not being an exception, has been regarded as one of the key notions in corpus linguistics. Generally, the criteria for classifying collocation are stability, frequency of use and lexical co-occurrence. However, defining collocation within the context remains delicate as well as most theoretical attempts at revealing how collocation really operates in a language. The starting point on which most researchers agree is the fact that collocation is driven by lexico-grammatical characteristics which are language-specific.

Key words. Collocation. Colligation. Mutual expectancy. Lexical stability. Formal stability. Frequency of use. Lexical co-occurrence. Lexico-grammatical reciprocity. Polylexicality. Phraseology.

Introduction

Le terme « *collocation* » (du lat. *collocare*), qui a pris de l'importance dans la lexicographie, dans la linguistique fonctionnelle et dans la linguistique de corpus, a été introduit par J. R. Firth. Ensuite ce sont J. Sinclair et M. K. Halliday, les disciples de Firth, qui ont poursuivi leurs recherches dans cette voie.

En français, la phraséologie collocationnelle prend ses sources notamment dans les travaux de F. J. Hausmann, I. A. Mel'čuk et U. Heid.

Dans la terminologie tchèque cette notion apparaît grâce à F. Čermák¹.

Même si le terme de « *collocation* » constitue l'une des notions clé de phraséologie, employée d'une manière large non seulement par les spécialistes du domaine, sa délimitation ne cesse d'être discutée.

Toutes les tentatives de définition n'aboutissant pour l'instant à aucune congruence ont en commun les traits définitoires du **figement**, de **fréquence d'utilisation** et de **voisinage**.

L'objectif de cet article est d'examiner les acquis disponibles et de proposer de nouvelles réflexions à propos de cette notion employée souvent comme un terme « passe-partout » pour les phénomènes d'arrangement avant tout lexical, peu saisissables et explicables surtout par les locuteurs non natifs.

¹ Čermák (2006: 11)

1. Le figement

Le figement ou la fixité est le trait majeur et décisif dont les linguistes se servent pour l'exploration du phénomène phraséologique de n'importe quelle langue en général. Le processus de figement est véhiculé par l'usage qui débarrasse une unité discursive de son statut d'une formation composée de mots isolés et il lui implique les propriétés d'une séquence non compositionnelle.

Le figement revêt deux types :

- formel : l'ordre des constituants d'une unité phraséologique reste en général inaltérable même si sa place dans le discours change. Il en résulte le fait que les unités phraséologiques fonctionnent d'une manière stable en tant que d'autres lexèmes.
- sémantique : la signification d'une unité phraséologique reste fixe et, dans la majorité des cas, ne change pas au cours du temps.

Même si le trait de figement est mentionné en premier lieu comme élémentaire, les linguistes s'en servent pour repérer au sein de ce phénomène-ci plusieurs degrés permettant une catégorisation plus fine de différents types des unités phraséologiques - le figement va progressant des unités lexicales aux énoncés phrastiques (proverbes, dictons) qui attestent un figement complet ne permettant aucune transformation qu'elle soit formelle ou sémantique.

En fonction de ceci, les « *collocations* » sont les combinaisons *qui ne perdent pas leur sens propre et premier qu'ils possèdent séparément (sens littéral) et dont la fonction est proprement référentielle* (Rey, 2002: 70).

Agnès Tutin et Francis Grossman dans leur article *Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif* » oppose ainsi les collocations « *à des expressions complètement figées comme pomme de terre ou nid d'ange, pour lesquelles le sens du tout est difficilement prédictible à partir du sens des parties : l'interprétation sémantique ne se fait pas de façon compositionnelle* (Tutin, 2002: 7).

Déjà Charles Bally dans son « *Traité de stylistique française* » (1909) propose le terme de « *séries phraséologiques* » qui se situent à la frontière entre le précodé et le libre : *Entre les extrêmes (groupements passagers et unités indécomposables) se placent des groupes intermédiaires appelés séries phraséologiques (p. ex. les séries d'intensité ou les périphrases verbales)* (Bally, 1909: 66). [...] *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 »* (*Ibid.*, p. 70).

Mel'čuk et ses collaborateurs inventent le terme de « *semi-phrasème* »² qui reflète le bien fondé des analyses décrites ci-dessus : « *semi* » indique que le figement est à mi-voix et le terme « *phrasème* » est monosémique.

2. La fréquence d'utilisation

Le processus de la formation d'une « *collocation* » est à dimension diachronique, c'est-à-dire son statut d'une unité fixe est approuvé avec le temps. Pour cette raison les « *collocations* » peuvent accuser certains aspects historiques voire archaïques d'ordre :

- lexical – un ou plusieurs constituants ne sont plus employés comme les mots indépendants en français moderne,
- syntaxique – c'est avant tout l'absence des prédéterminants, ensuite

² Mel'čuk—Clas—Polguère (1995: 42)

- les constructions elliptiques et l'ordre inhabituel des mots qui constituent les marques frappantes sur l'origine de longue date.

3. Le voisinage

Le principe de voisinage lié, voire figé, obéit aux règles de combinabilité et compatibilité qui se manifestent dans les discours par les combinaisons figées (ou la distribution figée dans la terminologie anglo-saxonne). I. A. Mel'čuk, linguiste s'occupant de la lexicologie combinatoire, opère dans son article élaboré en concours avec L. N. Iordanskaja, N. Arbatchewsky-Jumarie avec le terme de « *cooccurrence lexicale* » qu'il définit comme *la capacité des lexèmes de se combiner en syntagmes pour exprimer un sens donné* (Mel'čuk, 2007a: 3-34).

Ils distinguent les combinaisons libres de lexies résultant de « combinabilité facultative » et celles qui sont construites en fonction de contraintes conventionnelles – « *collocation* » définie comme *constituée d'une base, que le locuteur choisit librement en fonction de ce qu'il veut exprimer [...] et d'un collocatif [...], choisi pour exprimer un sens donné en fonction de la base* (Mel'čuk, 2007b: 20-21). Il est presque impossible d'élucider le processus qui lie un tel collocatif à une telle base. Pour la majeure partie de lexies, le locuteur n'a d'autre choix pour les acquérir que de les mémoriser.

Ce principe de « *cooccurrence lexicale* » qui est à l'origine de « *collocation* » a été désigné par Firth comme « *attraction mutuelle* »³ Il discerne en même temps la différence entre la « *collocation* » d'occurrences et la « *colligation* » dont l'usage est réservé aux relations mutuelles entre des catégories grammaticales (par exemple l'enchaînement des verbes avec leurs actants).

L'insertion de chaque unité lexicale dans la parole est régie d'une manière générale par les matrices inhérentes au système langagier. Ces tendances d'occuper une place régulière dans la chaîne parlée ont été analysées par exemple plus en détail par Michael Hoyer dans son œuvre *Lexical Priming : A New Theory of Words and Language* pour l'anglais.

Christopher Gledhill et Pierre Frath dans leur article « *Collocation; phrasème, dénomination : vers une théorie de la créativité phraséologique* » sont persuadés que *il serait approprié de remplacer l'opposition collocation / colligation par la notion plus synthétique de lexico-grammaire développée par Michael Halliday* (Gledhill, 2007: 3). Ils confirment ainsi le principe du fonctionnement synérgetique du signe lexical : *tout lexème fait partie d'un faisceau plus ou moins restreint de constructions lexico-grammaticales* (*Ibid.*, p. 4).

4. La polylexicalité

Certains linguistes essayent de définir les « *collocations* » aussi à l'aide des arguments d'ordre statistique. Par exemple Haussman caractérise les collocations comme les structures binaires, constituées de deux parties dont le statut est inégal et dont le rapport est dissymétrique.

Il s'agit du principe de polylexicalité qui est propre aussi à d'autres unités phraséologiques, car elles sont en général construites de plusieurs lexèmes qui peuvent être agglutinés ou séparés par les espaces blancs.

Ce postulat est contesté par I. G. Rey qui analysant la collocation « *marquer un but* » constate que « *certaines situations autorisent l'emploi de collocations réduites à un unique élément* » : « *But!* » (Rey, 2002: 95).

³ Firth (1957: 12)

Egalement F. Čermák faisant rappel à J. Sinclair confirme qu'en ce qui concerne le voisinage, on estime empiriquement qu'une « *collocation* » pour un mot donné, c.-t.-d. la combinaison primaire de formes verbales de texte, qui ne doivent pas être seulement binaires, apparaissent avec un rapport sémantique entre eux (et éventuellement avec une cohésion) jusqu'à la distance ± 5 mots – *Pokud jde o sousedství, empiricky se odhaduje, že kolokace pro dané slovo, tj. primárně kombinace slovních textových tvarů, které však nemusejí být jen dvoučlenné, se vyskytují se sémantickou vazbou mezi sebou (a popř. kohezi) až na vzdálenost ± 5 slov od něj, resp. od sebe* (Čermák, 2006: 11).

5. Typologie des collocations

Les typologies des « *collocations* » peuvent être relativement très divergentes en fonction des positions adoptées (fonctionnaliste, structuraliste, lexicographique, etc.) par les spécialistes.

Comme nous avons avisé ci-dessus que l'un des fondateurs de la phraséologie collocationnelle en français est F. J. Hausmann, il nous paraît approprié de commencer par sa propre typologie du phénomène collocationnel reproduite par I. G. Rey. En effet, ce linguiste opère avec deux taxinomies – l'une grammaticale et l'autre résultante des critères d'ordre sémantique.

La typologie grammaticale en sept structures (voir ci-dessous) pourrait être sans doute élargie sans prétendre à atteindre l'exhaustivité car le calcul statistique des combinaisons des mots autosémantiques avec des mots sysémantiques serait beaucoup plus élevé :

- N + adj. = *célibataire endurci*
- N sujet + vb. = [*chat*] *miauler*
- Vb. + N objet = *appeler l'ascenseur*
- N + prép. + N = *bouffée de chaleur*
- Vb. + prép. + N = *rougir de honte*
- Vb. + adv. = *désirer ardemment*
- Adv. + Adj. = *grièvement blessé*⁴

La distinction selon les aspects sémantiques relève trois catégories :

1. restreinte (*miauler [chat]*)
2. abstraite (*appeler l'ascenseur*)
3. figurée (*briser des chaussures neuves*)⁵

Agnès Tutin et Francis Grossman dans leur article « *Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif* » distinguent trois sortes de collocations sur le plan sémantique :

1. collocations « *opaques* »
2. collocations « *transparentes* »
3. collocations « *régulières* »⁶

Pour le premier type des collocations « *opaques* », l'association entre les constituants est d'ordre arbitraire et non transparente. Pour illustrer cette catégorie, ils mentionnent en tant qu'exemple « *peur bleue* ». Il est vrai que le lien entre les deux mots est opaque car il résulte de la transposition de sens de l'adjectif. Mais est-ce le trait qui appartient encore aux collocations ou bien déjà aux unités phraséologiques ? A notre avis, cette unité ne se retrouve

⁴ Rey (2002: 91)

⁵ Rey (2002: 91)

⁶ Tutin (2002: 22)

plus à la frontière entre le préconstruit et le libre ; son degré de figement est complètement accompli par la convention.

Les collocations « *transparentes* » sont celles qui ont « *un sens interprétable* » – (*grièvement blessé*)⁷, donc décodable à l'aide de leurs éléments constituants. Cette catégorie reflète le mieux le caractère de semi-figement si propre au phénomène collocatif.

Les collocations « *régulières* » sont dans la typologie de A. Tutin et F. Grossmann envisagées comme les associations motivées, transparentes est souvent à base unique comme en témoigne l'exemple convenablement choisi par les auteurs - « *nez aquilin* »⁸. Ce qui échappe un peu, c'est l'explication du bien fondé de l'adjectif, car la majeure partie de collocations peut être traitée comme régulière dans la mesure où le lien entre les éléments lexicaux est si fréquent qu'il peut être traité comme régulier, voire systématique.

La typologie moins complexe est proposée par *Isabel González Rey* dans son œuvre « *La Phraséologie du français* ».

En résumant les traits définitoires de « *collocations* » et en les comparant à d'autres unités phraséologiques, I. G. Rey a dressé un tableau où elle distingue les « *collocations* » tout simplement en fonction du sous-code à axe vertical et horizontal.

- « *en langue générale : mettre fin à qqch, avoir peur de qqch*
- *en langue de spécialité : (juridique) porter plainte, (commerciale) passer une commande* »⁹

Ce sont deux exemples de typologies des « *collocations* » qui peuvent être relativement divergente en fonction des positions (fonctionnaliste, structuraliste, lexicographique ; etc.) adoptées par les spécialistes.

Pour les buts pragmatiques nous nous permettons de proposer une typologie purement fonctionnelle basée sur les espèces de mots avec lesquels les « *collocations* » peuvent commuter : substantivales, adjectivales, verbales etc. Cette répartition suit en même temps la stratification de la langue selon trois axes : horizontal (les sous-codes de diverses activités humaines), vertical (le français soutenu, standard, familier, populaire etc.) et aréal (les dialectes, les variantes régionales, les patois).

Conclusion

Malgré les insuffisances dans les facteurs discriminatoires d'une « *collocation* » par rapport à d'autres unités phraséologiques hiérarchiquement superposées à celle-ci, nous postulons l'existence de ce type de signes, situés entre le mot et la phrase envisagés comme deux pôles extrêmes du discours.

Leur valeur énonciative d'un tout lexical découle de l'habitude que les locuteurs d'une communauté langagière relient un certain ensemble de mots pour référer à un objet de la réalité extralinguistique. Ces séquences lexicalisées sont constituées des unités lexicales qui coapparaissent avec un certain taux de fréquence sans pouvoir être substituées par les synonymes. Elles sont imposées par la langue en tant que norme.

Les « *collocations* » assument ainsi la fonction des référents ontologiques au même titre que d'autres unités lexicales. Elles augmentent également l'éventail de moyens d'expression qui ne sont dans aucune langue au monde limités aux mots simples et montrent comment il est fallacieux de vouloir donner un nombre exact pour telle ou telle langue.

⁷ Tutin (2002: 12)

⁸ *Ibid.* (2002: 12)

⁹ Rey (2002: 72)

Les entrées de n'importe quel dictionnaire incluent les « *collocations* » du mot base sans l'explication desquelles le lexique d'une langue naturelle serait considérablement réduit.

Résumé. Kolokace – pojem terminologicky nejednoznačný. Kolokace je jedním z klíčových pojmů korpusové lingvistiky stejně jako frazeologie jazyka. Obecně je tento termín definován pomocí kritérií ustálenosti, frekvence používání a lexikální sounáležitosti. Vymezení kolokace v rámci kontextu však přesto zůstává ošemetné, stejně jako teoretický výklad jejího fungování. Jediným společným bodem, na kterém se shoduje většina vědců je skutečnost, že realizace kolokace je řízena lexikálně-gramatickými vlastnostmi, jež jsou vlastní a typické pro každý jazyk zvlášť.

Bibliographie

- BALLY, Ch. (1909), *Traité de stylistique française*, Paris: C. Klincksieck.
- ČERMÁK, F., ŠULC, M. (2006), *Kolokace*, Praha: Nakladatelství Lidové noviny.
- FIRTH, J. R., (1957), *A synopsis of Linguistic Theory. 1930 – 1955*, Oxford: OUP.
- GLEDHILL, Ch., FRATH, P. (2007), “Collocation; phrasème, dénomination : vers une théorie de la créativité phraséologique”, In: *La Linguistique 43 (1)*, pp. 65 – 90.
- HOYE, M., (2005), *Lexical Priming : A New Theory of Words and Language*, London: Routledge.
- MELČUK, I. A., CLAS, I., POLGUÈRE, A. (1995), *Introduction à la lexicologie explicative et combinatoire*. Louvain: Duculot.
- MELČUK, I., IORDANSKAJA, L. N., ARBATCHEWSKY-JUMARIE, N. (2007), “ Un nouveau type de dictionnaire : Le dictionnaire explicatif et combinatoire du français contemporain”, In: *Cahiers de lexicologie N° 91, 2007-2*, pp. 3-34.
- MELČUK, I. A., POLGUÈRE, A. (2007), *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*, Bruxelles: De Boeck & Larcier s. a.
- REY, I. G. (2002), *La phraséologie du français*, Toulouse: Presses universitaires de Mirail.
- TUTIN, A., GROSSMANN, F. (2002), “Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif”, In: *Revue française de linguistique appliquée*, VII-1, pp. 7-25.

Jana Brňáková
Katedra romanistiky
Filozofická fakulta
Ostravská univerzita v Ostravě
Reální 5
CZ-701 03 Ostrava
République tchèque