

Multiword expressions in scientific discourse: a corpus-driven database

Agnès Tutin¹, Achille Falaise²

1 LIDILEM, Université de Grenoble Alpes
Université Stendhal Grenoble III, BP 25
38040 Grenoble cedex 9, France

2 LIG-GETALP, Université de Grenoble Alpes
385 avenue de la Bibliothèque,
38400 Saint Martin d'Hères, France

E-mail: agnes.tutin@u-grenoble3.fr; achille.falaise@imag.fr

Abstract

This paper presents an on-going project of corpus-cum-dictionary in the field of academic multiword expressions in French. This project is a lexicographic and pedagogical application of the Scientext corpus and aims to provide writing tools to help academic learners of French as a foreign language to produce high quality texts. Our system is based on a large syntactically annotated corpus and provides an onomasiological and a semasiological access to phraseological entries, including corpus examples stored in dynamic web pages.

Keywords: multiword expressions – collocations – attitudinal multiword expressions – corpora

1. Introduction

As outlined by several scholars, Loffler-Laurian (1995), Gledhill (2000), Pecman (2004), and Granger & Paquot (2010), fixed expressions are prevalent in the genre of academic and scientific writings and need to be fully mastered in order to become a member of the academic community. This paper describes an on-going project of a phraseological corpus-based database for French, where several kinds of multiword expressions can be accessed with several modes, mainly via a semasiological access and an onomasiological access. This project is a lexicographic and pedagogical application of the Scientext corpus, and aims to provide writing tools to help academic learners of French as a foreign language to produce high quality texts.

We will first describe the different types of academic multiword expressions addressed in this project. We then focus on attitudinal multiword expressions and their properties. Finally, the structure of our lexical database is presented as well as selected corpus examples.

2. Different kinds of academic multiword expressions

Our phraseological database includes frequent specific and cross-disciplinary multiword expressions pertaining to scientific activity, scientific reasoning, and scientific discourse (Cf Pecman 2004, Tutin 2007; Tutin forthcoming). The database includes several types of multiword expressions, each type being processed specifically within the database and the corpus. All the multiword expressions (MWE) are frequent and specific expressions were extracted from our corpora with NLP techniques. The different types of MWE are treated differently.

2.1 **Referential expressions** are widely studied in academic discourse. They are mainly lexical collocations (e.g. *faire une hypothèse* literally ‘make a hypothesis’) or frozen non-compositional expressions (e.g. *tenir compte* ‘to take into account’ or *point de vue* ‘point of view’) related to scientific activity, scientific objects or scientific reasoning. We posit, in line with Explanatory and Combinatorial Lexicology (Mel’čuk, 2011), that these two different kinds of MWEs should be differentiated. As collocations are compositional, each component will receive a semantic tag and be stored within the database entry (e.g. s.v. *hypothèse* for *refuter une hypothèse* ‘to disprove a hypothesis’) or *hypothèse valide* (‘valid hypothesis’). Fully frozen MWE expressions (e.g. *mettre en évidence* ‘highlight’) will be stored as single units.

2.2 **Discursive MWEs** aim at relating textual parts and are mainly structural and metatextual markers (*to begin with*, *in other words*) and logical connectives such as *that is why*. Mainly adverbials, but also conjunctions or prepositions, they are rarely specific to academic writing, in contrast to certain referential expressions. Nevertheless, many of them are more frequent in scientific writing such as *pour conclure* ‘to conclude’ (140 times more common than in literary and newspaper corpora) (Tran, forthcoming). Describing these MWEs requires explaining their discourse function (position, meaning and rhetorical function).

2.3 **Attitudinal MWEs** are mainly related to stance (e.g. *jusqu’à un certain point* ‘to a certain extent’) and modality (e.g. *il est nécessaire de* ‘it is essential to’). As previously found with discursive MWEs, the attitudinal MWEs discussed here are not fully specific of our genre, but occur more often in our corpora than in

general French. We selected the most frequent and the most specific attitudinal MWEs present in our academic corpora.

2.4 More complex **semantico-rhetorical routines**, fully specific of the scientific genre (see Sandor 2007; Teufel 1998; Tutin 2010; Granger & Paquot 2010; Jacques et al. 2013) are presented. They are associated to rhetorical functions such as comparison with peers or definition of the main research objective. In the Scientext project, which forms the source of our phraseological database, these semantico-rhetorical routines are extracted from the corpus with the help of local grammars by using syntactic relations.

3. Methodology

Our phraseological database is based on several principles:

1. All the phraseological resources are based on a large on-line corpus of scientific writings, Scientext, whose core is freely accessible via the web.
2. The database provides two kinds of access, an onomasiological access via semantic data, and a semasiological access via lexical entries. Simple semantic syntactic and certain discursive information is provided as well as information on frequency.
3. As the database is intended to be a writing aid and can be considered as a dictionary-cum-corpus project (Hartmann 2005; Granger & Paquot 2010), phraseological data are associated to selected on-line examples. A dynamic access to web pages is provided, with the possibility to access larger contexts and statistics, e.g. parts of texts where a multiword expression is most frequent.

Our database is constructed with the help of the Scientext corpus, syntactically annotated with the Syntex parser (Bourigault 2007). Syntactic relations between words are particularly useful to extract complex MWEs in which the components may be non-contiguous. An example of a parsed sentence is given in Figure 1.

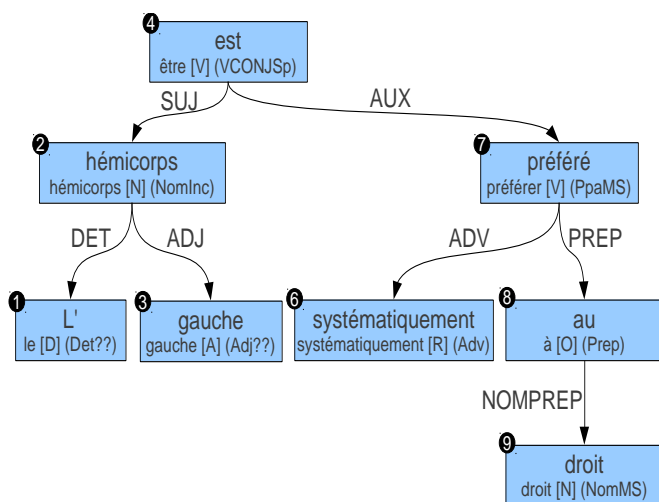


Figure 1 : Syntactic analysis of the sentence *L'hémicorps gauche est préféré systématiquement au droit*

The construction of the database follows several stages:

1. **Stage 1: Extraction of frequent and cross disciplinary MWEs** with simple NLP techniques, mainly n-gram techniques and lexical associations. Other techniques using lexico-syntactic patterns (see Corman 2012) will be experimented in this project.
2. **Stage 2: Study and linguistic analysis of MWEs, based on corpus examples.** For this analysis, examples are manually extracted with Scientext queries. Some queries may be quite complex when local grammars in which grammatical relations are involved (Falaise *et al.* 2012). This is especially the case with semantico-rhetorical patterns (Tutin 2010), as in the following example on “formulating a hypothesis”.

```
//TITRE : Formulation d'une hypothèse

//INFO: Expressions pour formuler une hypothèse

(OBJPASSIF, #3, #1) =
(SUJ, #1, #2) (AUX, #3, #2)
$formul=avancer, émettre, effectuer, faire
, formuler, poser, prendre, proposer, retenir, soutenir, utiliser
Main = <lemma=$formul, #1> &&
<lemma=hypothèse, #2>:: (OBJ, #1, #2) OR
(OBJPASSIF, #1, #2) OR (ADJ, #2, #1)
```

Figure 2: An example of a local grammar associated with the semantico-rhetorical routine “formulating a hypothesis”

Here, the query indicates that several verbs such as *formuler* (‘formulate’) or *émettre* (‘emit’) may have a direct object *hypothèse*. Passive alternation and past participle passive constructions are also included with the help of syntactic relations.

Figure 3 shows concordances associated to the query illustrated in Figure 2. It is interesting to note that syntactic relations enable one to extract word associations that are separated by a large number of words as in example 5, where *proposer* and *hypothèse* are separated by 15 words) (*l'hypothèse selon laquelle la multiplication de l'ADN pourrait participer au contrôle de l'expression des gènes est proposée*). Syntactic analysis is essential to reducing noise and silence in queries, i.e. inaccurate examples in corpora or examples which would not be extracted by simpler techniques.

<input checked="" type="checkbox"/>	1	L'	hypothèse émise	pour explorer les variations dans les usages adultes est la
<input checked="" type="checkbox"/>	2	point de vue de sa stratification sociale on peut	émettre l' hypothèse	que les parents donnent la priorité à l' apprentissage de
<input checked="" type="checkbox"/>	3	Von Hippel , chercheur de longue date dans le domaine de l' innovation économique ,	fait l' hypothèse	que nous sommes entrés dans une ère de " démocratisation
<input checked="" type="checkbox"/>	4	de régulation de l' activité des gènes , et	formulent l' hypothèse	selon laquelle des mécanismes analogues expliquent les variations d' activité
<input checked="" type="checkbox"/>	5	bactéries , des modifications possibles de l' ADN , l'	hypothèse selon laquelle la méthylation de l' ADN pourrait participer au contrôle de l' expression des gènes est proposée	simultanément par Arthur Riggs 5 et Robin Holliday 6 en 1975
<input checked="" type="checkbox"/>	6	Ils	émettent l' hypothèse	qu' elle soit due à une accumulation de structures tourbillonnaires
<input checked="" type="checkbox"/>	7	pas tous aux mêmes attributs de la dissertation et nous	faisons l' hypothèse	que cette ambiguïté engendre le risque d' enseigner une vision
<input checked="" type="checkbox"/>	8	Ainsi , comme nous en	faisons l' hypothèse	, le sens de la notion de " problématique "

Figure 3: Concordances associated to the semantico-rhetorical routine “formulating a hypothesis”

Some queries are simpler, since MWEs can already be encoded in the NLP system as lexical units, for example, the the simple query to extract of the MWE *à juste titre* (‘rightly’) is shown in Figure 4.

Mot 1
Lemme à juste titre

Recherche

Aperçu des résultats

19 occurrences trouvées

N°	Contexte gauche: mots	Occurrence:	Contexte droit: mots
<input checked="" type="checkbox"/>	1 C. Fuchs insiste par ailleurs	à juste titre	sur le caractère formateur et l' intérêt didactique de la
<input checked="" type="checkbox"/>	2 faire retomber dans ses " motivations extérieures " que dénonce	à juste titre	B. Lahire .
<input checked="" type="checkbox"/>	3 la lecture " , parue en novembre 2000 , remarque d'ailleurs	à juste titre	le peu d' études qui ont été consacrées au niveau morphématique dans
<input checked="" type="checkbox"/>	4 57) souligne	à juste titre	que la restauration de la dimension argumentative dans le travail
<input checked="" type="checkbox"/>	5 , la lecture fait l' objet à l' école ,	à juste titre	, de tout un effort visant à la promouvoir ,
<input checked="" type="checkbox"/>	6 Comme l' observe	à juste titre	un économiste contemporain , " notre appareil productif et notre
<input checked="" type="checkbox"/>	7 secondaires , connus sous le nom de virus informatique x	À juste titre	, ces virus n' ont pas manqué de susciter
<input checked="" type="checkbox"/>	8 pas manqué de susciter) (with good reason ,	À juste titre) (alarm and horror , panique) (in

Figure 4: An example of query for MWEs encoded as lexical units

Once MWEs are extracted, they are analyzed and encoded in a lexical database with several kinds of information: semantic, syntactic, discursive, or related to frequency. Detail examples of attitudinal MWEs will be discussed in Section 4.

3. Stage 3: Selection of relevant on line concordances associated with MWE.

Concordances and authentic examples are invaluable for pedagogical purposes. NLP corpus based projects such as Alfalex are very promising in this perspective (Verlinde 2009). Nevertheless, annotated corpora with the help of NLP techniques cannot be used without human intervention mainly for two reasons. First of all, NLP techniques are not error-free and some extracted examples can be erroneously analyzed, for example an incorrect part-of-speech category for a word. Secondly, some examples may appear inappropriate because they are too complicated to understand, especially for non-native speakers, or they may be atypical. For this reason, we have developed a system to filter on-line concordances, for which larger contexts and statistics remain available. Figure 5 demonstrates how to save the appropriate examples in a file, which will be uploaded and associated to an MWE entry. In this example, we have only selected occurrences for the MWE *à l'évidence* (‘obviously’, ‘clearly’). Examples including the other MWE *se rendre à l'évidence* (‘face the facts’) have been discarded.

6 Il faut se rendre à l'évidence .

7 Si la violence comportementale s' impose à l'évidence , il n' en est pas de même d' autres

8 Il faut se rendre à l'évidence , il y a loin de la théorie écrite à la

9 négativement par les éléments cotextuels , ce qui , à l'évidence , n' est pas le cas ici .

10 dans un rapport potentiel de substitution là où il manque de personnel face à des cohortes importantes d' étudiants à former

11 Pourquoi les gens ne parlent -ils pas comme ils estiment à l'évidence qu' ils le devrait ? " .

12 Par contre , leur durée de vie est à l'évidence plus courte .

Exporter les concordances au format HTML

Exporter les concordances au format ODS (LibreOffice)

Exporter les concordances au format XLS (Excel)

Exporter les concordances au format Gnumeric

Exporter les concordances au format CSV

Exporter les concordances au format SQR

Figure 5: On line selection of concordances (for reuse in the MWE database)

The example file can be saved and associated to a MWE as discussed in the following section.

4. Structure of the database: the example of attitudinal MWEs

MWE entries are integrated into an (online) database, included in several kinds of sub-databases, depending on the kind of MWE expressions: discursive MWEs are not encoded in the same manner as semantico-rhetorical routines. The structure of the MWE database is quite rudimentary, since more attention is paid to corpus examples, in the perspective of a writing aid and for

CALL activities. Databases on discursive MWEs and attitudinal MWEs are under development (for discursive MWEs, see Tran forthcoming). Some semantico-rhetorical routines have already been developed (See Tutin 2010, Jacques *et al.* 2013 for English) but they need to be extended.

4.1 Fields of the database Our linguistic description includes several fields for MWEs. It is intended both for pedagogical purposes, mainly as a writing aid in French as a Foreign Language, and for NLP and corpus linguistics extractions.

Our lexical entries include several fields, the main fields are presented with examples in Figure 7.

1. Lexical entries and variants of lexical entries. For example, *de notre point de vue* ('in our opinion') has a pronominal variant *de mon point de vue* ('in my opinion').
2. Semantic types are associated to MWEs, especially in the perspective of an onomasiological access. Several semantic types are given such as: 'certainty/evidence', 'opinion', 'positive opinion', 'negative opinion', 'approximation', 'intensity', 'necessity', 'probability'.
3. Part of speech of the MWE. A POS is given for the whole MWE in order to fully grasp the syntactic behavior of the MWE. Most attitudinal MWEs are adverbials but some of them are more complex (e.g. *il est possible*, *il est probable*).
4. Part of speech of the components, e.g. for *pour ainsi dire* ('so to speak'). This information is useful for NLP applications and corpus extraction.
5. Frequency. This information is highly relevant for pedagogical purposes. We indicate an approximate value (***: frequent, **: quite frequent, *: not so frequent). A precise value is given in the database corresponding to the number of appropriate examples extracted from the corpus.
6. Distribution of adverbials within the sentence is a very complex issue for non-native speakers. A special attention is paid to this issue. For example, *à première vue* ('at first sight') can appear, as can be seen in Figure 6, at the beginning of the sentence (example 1), before the verb (example 6) or more frequently just after the verb (examples 2,3,4,5).

1.	On en voit un nouvel exemple ici : "	à première vue	ces analogies pourraient venir nous aider dans notre tâche
2.	être induit par des agents chimiques n'ayant ,	à première vue	, aucun dénominateur commun : métaux de transition ,
3.	ce remboursement à l' Opérateur du système qui peut	à première vue	s' apparenter à un coût pour le producteur G1
4.	Il semble donc	à première vue	que l' extraction directe des coefficients soit impossible
5.	permettraient de rendre compte d' un écart figural paraissent	à première vue	extrêmement ténus .

6.	U. Eco Eco (U.) ,	à première vue	, ne s' intéresse qu' à cette dernière distinction
----	---------------------	-----------------------	--

Figure 6: Examples of the distribution of *à première vue*

7. Synonyms are introduced, preferably MWEs.
8. Short definitions are provided.
9. A link to concordances is given.

expression	type	Cat.	Freq .	position dans la phrase	synonymes	Définition
à première vue	la certitude//l' évidence	adv	*	préverbale postverbale tête de phrase		au premier examen, apparemment
bien entendu	la certitude//l' évidence	adv	**	postverbale tête de phrase préadjectivale	bien sûr	de façon évidente
bien sûr	la certitude//l' évidence	adv	***	postverbale tête de phrase	bien entendu	de façon évidente, avec certitude

Figure 7: Main fields of the database of attitudinal MWEs

4.2 On-line Database

The phraseological database is integrated into a simple on-line dictionary-cum-corpus application. Our application is intended to be more of a corpus-cum-dictionary application than its opposite, since we believe that learners need most is a large number of appropriate examples. The dictionary is nevertheless useful to easily access the corpus information.

The phraseological entries can be accessed through an onomasiological access, with the help of the semantic types presented above. We are currently incorporating data for attitudinal MWEs and discourse MWEs as well as semantico-rhetorical routines, but our final application will also include referential MWEs, be they collocations or frozen expressions¹. Figure 8 provides four MWE expressions, sorted by descending frequency, of onomasiological access with the help of the semantic type "approximation". The user has chosen one of these expressions *au sens large* ('in the wide sense') and a brief lexical entry is provided for the word including the POS, a definition and synonyms (with links in the database when available). Selected examples are also given on concordances, which can be dynamically extended to a larger context (Figure 9). Statistics (on genres – research articles, communications or theses –, textual sections and disciplines) are also available.

¹ These MWEs are far more numerous than those previously mentioned and collocations, being compositional, are complicated to implement.

Expressions indiquant le point de vue et l'attitude dans les écrits scientifiques

Accès par sens ▼

Pour exprimer...

Expressions :	Catégorie
... la certitude/l'évidence	
... l'approximation	adverbe
... l'opinion positive	
... le point de vue	
... la quantité/ l'intensité	
... l'opinion négative	
... la qualité	

Expressions :	Définition	Synonymes
en général (***)		
en moyenne (**)		
au sens large (*)	en ayant une définition extensive, large des phénomènes	
pour ainsi dire (*)		grosso modo

Exemples

N°	Contexte gauche: 10 mots	Occurrence:	Contexte droit: 10 mots
1	lecteur , nous allons exploiter les différentes connaissances linguistiques (au sens large	du terme) présentes dans les textes .
2	Je vais vous parler de ma formation	au sens large	, et de mes expériences dans l'enseignement et la
3	est question de tenir compte de tous les textes (pris ici	au sens large	puisqu' il s' agit de caricatures , mais nous verrons
4	Il s' agirait davantage de la situation ou	au sens large	, du contexte culturel et historique dans lequel s' inscrit

Figure 8: An example of onomasiological access: MWEs of approximation

Pour exprimer...

Expressions :	Catégorie
... la certitude/l'évidence	
... l'approximation	adverbe
... l'opinion positive	
... le point de vue	
... la quantité/ l'intensité	
... l'opinion négative	
... la qualité	

Expressions :	Définition	Synonymes
en général (***)		
en moyenne (**)		
au sens large (*)	en ayant une définition extensive, large des phénomènes	
pour ainsi dire (*)		grosso modo

Exemples

1	lecteur , nous allons exploiter les différentes connaissances linguistiques (au sens large du terme) présentes dans les textes .
---	--

[tal-art-texte_thèse_scientext_239-body] - Navigation textuelle : représentation des textes et des connaissances (Traitement Automatique des Langues, vol 47, n°2)

que la seconde concerne les liens qui sont présentés à l'utilisateur.

Tandis que l'hypertexte adaptatif conserve la notion de lien tel qu'il existe dans l'approche traditionnelle de l'hypertexte, l'hypertexte dynamique change la notion même de lien. Dans cette approche, les liens sont créés à l'exécution, au lieu d'être calculé préalablement (liens pré-calculés), ou à travers des modifications ou sélections dans des ensembles existants de liens (liens adaptatifs) (Bodner et Chignell, 1999). Une possibilité dans cette approche consiste à calculer les liens selon des relations prédéfinies ou selon des critères de similarité entre les textes. Dans ce cas, un lien n'est pas défini comme un pointeur d'un noeud hypertexte vers un autre, mais comme une requête visant un noeud. Ces requêtes peuvent prendre en considération l'historique de navigation, le profil de l'utilisateur, etc.

Cette notion de requête sur un noeud est une source d'inspiration pour notre démarche, mais plutôt que de considérer des éléments liés au lecteur, nous allons exploiter les différentes connaissances linguistiques (au sens large du terme) présentes dans les textes. Cette exploitation exige une étape d'annotation du texte, résultat d'un traitement linguistique finalisé.

Figure 9: Larger contexts for *au sens large*

Beside onomasiological access, semasiological access is also available (Figure 10). MWEs are here sorted by alphabetical order.

Expressions indiquant le point de vue et l'attitude dans les écrits scientifiques

Accès par expression ▼

Expressions :

à l'évidence OK

Expressions :	Catégorie
bien sûr (***)	adverbe
sans doute (***)	
en général (***)	
bien entendu (**)	
en moyenne (**)	de façon irréfutable
à bon escient (*)	
à juste titre (*)	
à l'évidence (*)	de toute évidence, sans conteste

Exemples

N°	Contexte gauche: 10 mots	Occurrence:	Contexte droit: 10 mots
1	déroulement de la situation dans la classe montre qu'	à l'évidence	, la dévolution de la situation n' a pu avoir lieu
2	Le tableau ci-dessous montre	à l'évidence	une modification importante des valeurs prises par les variables ,
3	des devoirs d' élèves de ces mêmes classes , montrent	à l'évidence	les mêmes problèmes de contrôle du sens des opérations successives

Figure 10: An example of semasiological access: the example of *à l'évidence*

5. Perspective and conclusion

Our project can be considered more as a corpus-cum-dictionary than a dictionary-cum-corpus project. It aims at showing and teaching academic phraseology to non-native speakers of French with the help of authentic selected examples. The lexical database can be considered essentially as an onomasiological access to MWE occurrences drawn from the corpus. The originality of our database is that it will largely be built on a syntactically annotated corpus both for the extraction of MWEs and for the selection of relevant examples. Dynamically selected on-line concordances enable the learner to explore phraseological units in authentic scientific texts.

Several extensions are being considered. As in the LEAD Project or the BLF project (Verlinde 2009), we plan to associate CALL activities with this database by using annotated corpora including phraseological and semantic information on cross-disciplinary phraseological lexicon. First, we plan to display different kinds of MWE annotations, e.g. structural MWEs and attitudinal MWEs on the same text. Freely available research articles of the Scientext corpus could be used for this purpose. Second, we plan to build on-line exercises on specific aspects of MWEs, following several studies conducted by our research team in French (Cavalla & Loiseau, forthcoming) and in English (Hartwell & Jacques, 2012).

6. Acknowledgements

Special thanks to Laura Hartwell and Patricia Andreoli-Jones who very kindly read and commented a previous version of this paper.

7. References

- Cavalla C., Loiseau M. (forthcoming). Scientext comme corpus pour l'enseignement. In Grossmann F. ; Tutin A. (eds.), *L'écrit scientifique : du lexique au discours : Autour de Scientext*, Rennes, PUR.
- Falaise, A., Tutin, A., Kraif, O. (2012). Une interface pour l'exploitation de corpus arborés par des non informaticiens : la plate-forme ScienQuest du projet Scientext, *Traitement Automatique des Langues*, Volume 52 (n° 3), pp. 201-228.
- Gledhill C. (2000), *Collocations in science writing*, Language in performance, 22, Tübingen, Gunter.
- Granger, S., Paquot, M., (2010). The Louvain EAP Dictionary (LEAD) », *Proceedings of the XIV EURALEX International Congress*, Leeuwarden (The Netherlands), 6-10 July 2010, pp. 321-326.
- Hartmann R.K.K. (2005), Pure or hybrid? The development of mixed dictionary genres, *Facta Universitatis* 3(2), pp. 193-208.
- Hartwell, L., Jacques M.-P. (2012), A corpus-informed text reconstruction resource for learning about the language of scientific abstracts, *Proceedings of the International EuroCALL Conference*, University of Gothenburg (Suede), 22-25 august 2012, pp. 117-123.
- Jacques, M.-P., Hartwell, L., Falaise, A. (2013). Techniques de TAL et corpus pour faciliter les formulations en anglais scientifique écrit, *Actes de TALN*, 17-21 juin 2013, Les sables d'Olonne.
- Kosem, I. (2010). Designing a model for a corpus-driven dictionary of Academic English. PhD thesis. Aston University, Birmingham, UK.
- Krishnamurthy, R. (2008). Corpus-driven lexicography. *International Journal of Lexicography*, 21(3), pp. 231-242.
- Loffler-Laurian A.-M. (1995), « Locutions et discours scientifiques », Martins-Baltar M. (dir.), *Cahiers du Français Contemporain*, La locution en discours ; n° 2, 243-269. *Macmillan English Dictionary Online*. Accessed at: <http://www.macmillandictionary.com>.
- Pecman M. (2004), *Phraséologie contrastive anglais-français : analyse et traitement en vue de l'aide à la rédaction scientifique*, Thèse de doctorat, Université de Nice Sophia Antipolis.
- Sándor A. (2007), Modeling metadiscourse conveying the author's rhetorical strategy in biomedical research abstracts, *Revue Française de Linguistique Appliquée*, XII, 2007-2, pp. 97-108.
- Teufel, S. (1998), Meta-discourse markers and problem-structuring in scientific articles", *ACL 1998 Workshop, Discourse Structure and Discourse Markers*. Montreal, Somerset/New Jersey, pp. 43-49.
- Tutin, A. (2010). Showing phraseology in context: an onomasiological access to lexico-grammatical patterns in corpora of French scientific writings. In Granger S., Paquot M. (eds.), *eLexicography in the 21st century: New challenges, new applications, Proceedings of eLex 2009, Louvain-la-Neuve, 22-24 October 2009*. Louvain-la-Neuve: Presses Universitaires de Louva, pp. 302-312.
- Verlinde, S. (2009). The Base lexicale du français: a multi-purpose lexicographic tool. In S. Granger, M. Paquot (eds.) *eLexicography in the 21st century: New challenges, new applications, Proceedings of eLex 2009, Louvain-la-Neuve, 22-24 October 2009*. Louvain-la-Neuve: Presses Universitaires de Louvain, pp. 335-342.