Lexicon Embedded Syntax

Alain Polguère
ATILF UMR 7118, CNRS-Université de Lorraine
44 avenue de la Libération, BP 30687
54063 Nancy cedex, France
alain.polguere@univ-lorraine.fr

Abstract

This paper explores the notion of lexicon embedded syntax: syntactic structures that are preassembled in natural language lexicons. Section 1 proposes a lexicological perspective on (dependency) syntax: first, it deals with the well-known problem of lexicon-grammar dichotomy, then introduces the notion of lexicon embedded syntax and, finally, presents the lexical models this discussion is based on: lexical systems, as implemented in the English and French Lexical Networks. Two cases of lexicon embedded syntax are then treated: the syntax of idioms, section 2, and the syntax of collocations, section 3. Section 4 concludes on the possible exploitation of syntactic structures that can be extracted from lexical systems.

1 Lexicological Perspective on Syntax

1.1 Lexicon-Grammar Dichotomy

The task of modeling languages is often equated with a task of writing so-called *grammars*. This is clearly demonstrated by the fact that most theoretical proposals in modern linguistics are designated as specific types of grammars: Generative Grammar, Case Grammar, Lexical Functional Grammar, Word Grammar, Generalized Phrase Structure Grammar, Construction Grammar(s), Role and Reference Grammar, Functional Discourse Grammar, etc. (Polguère, 2011, pp 82–83). It should be noted that this focalization on an all-encompassing notion of grammar runs deep. For instance, the 1795 law that created the school of oriental language studies in France (INALCO[1]) specified as follows the linguistic descriptive task assigned to its professors:

> "Lesdits professeurs composeront en français la grammaire des langues qu'ils enseigneront: ces divers ouvrages seront remis au comité d'instruction publique."[2]

No mention of a need to compile **dictionaries** for oriental languages, as if it were natural to designate with the term *grammar* the main tool to be used by XVIIIth century officials and merchants for communicating with "locals". It should be stressed that this rather confusing notion of Grammar – with a capital *G* – is extremely broad and encompasses the set of **all** linguistic rules that make up a natural language. It is distinct from the grammar as a language module that stands in opposition with its functional counterpart: the lexicon. Both linguistic modules have been loosely characterized as follows by O. Jespersen – in terms of their corresponding fields of study:

> "[g]rammar deals with the general facts of language, and lexicology with special facts" (Jespersen, 1924, p 32).

In the present discussion, we will strictly abide by the above characterization and consider the grammar of a language as being the system of all general rules of that language – i.e. rules that are not properties assigned to given words – and the lexicon of that language as being the system of all its word-specific rules.

It is a well-established fact that there exists a blurry demarcation between grammar and lexicon (Keizer, 2007). Rules that are specific to linguistic entities that present analogies with "words" but are not strictly speaking lexical units are less lexical in nature and possess a certain grammatical flavor. For instance, rules that account for the properties

[1] http://www.inalco.fr

[2] 'Said professors will elaborate in French the grammar of languages they will be teaching: these various books will be submitted to the public instruction committee.'

Proceedings of the Third International Conference on Dependency Linguistics (Depling 2015), pages 2–9,
Uppsala, Sweden, August 24–26 2015.

of bound morphemes (the English derivative suffix *-ly*, the prefix *poly-*, etc.) belong to the lexicon because they are specific to a linguistic sign, hence not general, but they are borderline due to the morphological nature of the sign in question. In what follows, quite a few linguistic entities will be presented as belonging to lexical models based on this preliminary characterization of the respective scope of grammar and lexicon and in spite of widespread practices that may tend to view lexicons strictly as repositories of lexical units.

1.2 Focus on Lexicon Embedded Syntax

Another factor that blurs the lexicon-grammar partition is the very fact that, in any natural language, a considerable number of syntactic structures are preassembled in the lexicon. Valency-controlled dependencies – whose modeling is directly relevant to lexicological studies – are the most obvious manifestation of this phenomenon. A valency dictionary or lexical database (Fillmore et al., 2003; Mertens, 2010) is nothing but a lexicographic description of a significant part of lexicon embedded syntax. This fact is now widely acknowledged. What is much less known and/or taken into account, specially in Natural Language Processing, is the extent to which syntactic structures of natural languages find their origins in lexicons, thanks to the omnipresence of phraseology (Becker, 1975).

In what follows, we will focus of two types of lexicon embedded syntactic structures:

- lexico-syntactic structures of idioms (section 2);
- collocational syntactic structures (section 3).

We are particularly interested in showing how a rich formal lexical model (see 1.3 below) can account for lexicon embedded syntax and serve as repository of "canned" syntactic structures that are directly extractable from lexical data.

1.3 Lexical Systems

In order to provide data for the proper treatment of lexicon embedded syntax, lexical models need to have "phraseological genes": they have to be based on theoretical and descriptive principles that fully take into consideration the omnipresence of phraseology in natural languages. Such is the case of Explanatory Combinatorial Lexicology (Mel'čuk et al., 1995; Mel'čuk, 2006), that is being used as theoretical background in the present discussion. More specifically, we will refer to a new type of lexical model built within this framework – lexical systems (Polguère, 2009) –, using two specific instances of such models: the *English* and *French Lexical Networks* – hereafter, *en-* and *fr-LNs*.

Lexical systems are huge graphs of interconnected lexical entities. Polguère (2014) discusses the rationale behind the choice of this particular type of structure, formally characterized by four main properties.

Property 1. The lexical system of a language $\mathscr{L}$ is mathematically defined as an oriented graph: a set of nodes and a set of oriented edges (= ordered pairs of nodes).

- Nodes correspond, first, to lexical units of $\mathscr{L}$ (lexemes and idioms) and, second, to quasi-lexical units (linguistic clichés, proverbial clauses, etc.).
- Edges correspond primarily to Meaning-Text lexical function relations (Mel'čuk, 1996).[3]

Property 2. Nodes of the graph are non-atomic entities. They are "containers" for a rich variety of semantic and combinatorial information about the corresponding unit (grammatical characteristics, definition, etc.); they also contain pointers to lexicographic examples (sense illustrations), their content being informationally analogous to that of dictionary articles (Polguère, 2014, pp 15–16).

Property 3. Lexical systems possess a non-ontological graph structure that belongs to the family of so-called *small-world networks*. As such, they display remarkable mathematical properties (Gader et al., 2014, §3) that can be used to extract node clusters corresponding to *semantic spaces* (Polguère, 2014, §2.2.2).

Property 4. Each important piece of information in lexical systems (existence of a lexical unit, assignment of a grammatical characteristic, lexical link, etc.) possesses an associated measure of

[3] Other relations are, at the moment: copolysemy links (FOREST **1** [*of oak trees*] and FOREST **2** [*of antennas*] belong to the same polysemic vocable and are connected by a relation of metaphor), definitional inclusions (the meaning of DOG is included in the definition of [to] BARK) and formal inclusions (the lexeme BULLET is formally included in the lexico-syntactic structure of the idiom BITE THE BULLET) – we will examine this latter type of relation in section 2 below.

confidence that can be used to perform probabilistic computing on the graph. Measurement of confidence is particularly relevant for the implementation of *analogical reasoning* on lexical models.

Figure 1 illustrates the graph structure of lexical systems. It visualizes a semantic space controlled by the French lexeme FORÊT 'forest' in the fr-LN. In this figure, spatialization and coloring of nodes visualize the result of an automatic semantic clustering performed on the lexical graph; this mode of visualization reflects semantic proximity inferred from the topology of the graph (Chudy et al., 2013).

Work on lexical systems started with experiments on the mechanical compilation of traditional Explanatory and Combinatorial models (Polguère, 2009), then evolved into full-scale lexicography with the construction of the fr-LN, the first manually-built lexical system (Lux-Pogodalla and Polguère, 2011; Gader et al., 2012). While lexicographically developing the fr-LN, a first version of a lexical system for the English language – the en-LN – has been automatically compiled from the Princeton WordNet (Gader et al., 2014). This latter lexical system offers a large-scale coverage of English in terms of wordlist. It is however essentially based on synonymy-like relations, inherited from WordNet; only the fr-LN fully reflects the amplitude of both paradigmatic and syntagmatic lexical function relations. Additionally, it is only in the fr-LN that the actual Explanatory Combinatorial approach to phraseology is fully implemented at present. For this reason, we will need to use both French and English illustrations in the following discussion, depending on the availability of data in the current language models.

Table 1 gives statistics on the en- and fr-LNs in their present state.

Graph characteriscs	en-LN	fr-LN
Num. lexical units = senses (LU)	206 995	26 020
Num. vocables = dict. entries (V)	156 587	16 981
Polysemy rate (LU/V)	1.32	1.53
Num. lexical functions links (LFL)	945 971	49 539
Num. other links (OL)	46	13 672
Connectivity rate ((LFL+OL)/LU)	4.57	2.43

Table 1: Current statistics on the en- and fr-LNs

2 Syntax of Idioms

We can now proceed with the examination of the first type of lexicon embedded syntax: the syntax of idioms. By this we mean lexico-syntactic structures that are associated with idioms in the fr-LN.[4]

Because they are semantically non-compositional, idioms are considered as full-fledged lexical units in Explanatory Combinatorial Lexicology. For this reason, they possess, just like lexemes, their own individual description in the fr-LN.

On the one hand, the behavior of idioms is known to be highly irregular (for instance, some idioms allow syntactic modification on some of their lexical constituents and other do not); on the other hand, it can be expected that general rules could be identified that condition part of idioms' behaviors, based on their lexico-syntactic structure. For this reason, it has been decided to specify, for each individual idiom in the fr-LN wordlist, its constitutive lexemes and its basic syntactic structure (Pausé, to appear). This is implemented as follows.

First, each phrasal part of speech – nominal idiom, verbal idiom, etc. – is linked to a set of syntactic templates that identify possible syntactic structures for idioms belonging to this part of speech. For instance, the *verbal idiom* part of speech (Fr. *locution verbale*) is associated, among others, with a syntactic template named `V Art NC` ('Verb + Article + Common noun') that designates the syntactic structure shown in Figure 2.

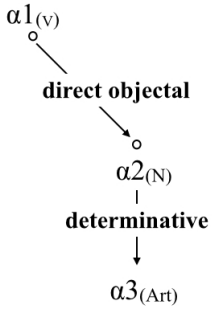

Figure 2: Syntactic structure of the `V Art NC` idiom template.

Second, each time an idiom is created in the fr-LN, two operations are performed:

1. the newly created idiom is linked to one of the syntactic templates associated to its part of speech;

[4]Work on assigning lexico-syntactic structures to idioms in the en-LN has not started yet and all our examples in this section will therefore be borrowed from French.

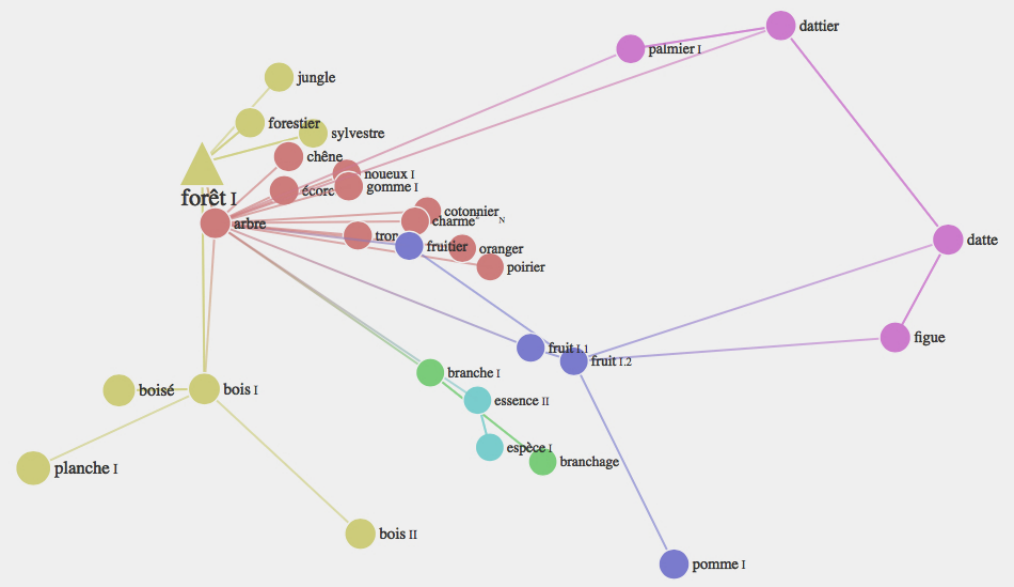

Figure 1: Semantic space controlled by Fr. FORÊT I 'forest' in the French Lexical Network (fr-LN)

2. lexical nodes in this syntactic template are linked to actual lexical units that make up the idiom.

For instance, Figure 3 shows how the lexico-syntactic structure of the idiom ⌜SUCRER LES FRAISES I⌝ 'to tremble because of advanced age' (lit. 'to sugar.the.strawberries')[5] is specified on the `V Art NC` template using the fr-LN lexicographic editor. In this figure, names appearing in the `Sense` column correspond to actual pointers to lexemes (senses) of the fr-LN; names in the `Form` column are only wordforms that will be used when displaying the instantiated syntactic template. (If nothing is specified, the name in the corresponding `Sense` cell will be displayed.)

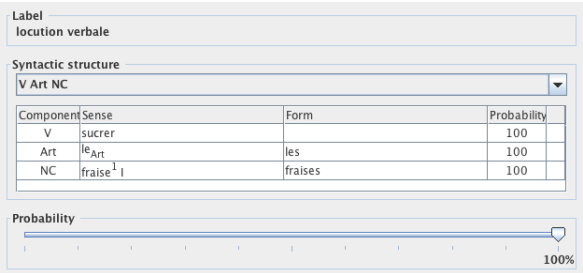

Figure 3: Specifying a lexico-syntactic structure.

Once the lexico-syntactic structure of ⌜SUCRER LES FRAISES I⌝ has been fully instanciated (Figure 3), in can be interpreted by the general – hence, grammatical – syntactic template of Figure 2 in order to derive the fully lexicalized syntactic structure shown in Figure 4.[6]

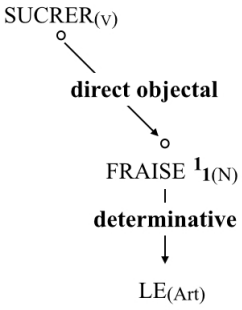

Figure 4: Syntax of ⌜SUCRER LES FRAISES I⌝.

To our knowledge, the fr-LN is the first lexical database that systematically accounts for the lexico-syntactic structure of idioms it contains – in point of fact, current lexical resources seldom provide individual descriptions for idioms. At present, it is possible to derive from fr-LN data 3,018 syntactic structures of individual idioms (such as that in Figure 4), which is only a small portion of the syntax of idioms embedded in the French lexicon.

[5] There is another sense ⌜SUCRER LES FRAISES II⌝, derived from the first one, that means 'to be senile'.

[6] An important piece of information is missing in this structure: the fact that the lexeme FRAISE $^{1}_{1}$ has to carry the grammeme 'plural' (⌜*sucrer les fraises*⌝ and not *⌜*sucrer la fraise*⌝). The fr-LN does not support yet the specification of grammemes in idiom syntactic structures.

3 Syntax of Collocations

3.1 Functional notion of collocation

We now examined a second case of lexicon embedded syntax: the syntax of collocations. *Collocation* is understood here as designating a functional rather than statistical notion (Hausmann, 1979); it can be defined as follows.

> A collocation, e.g. *to run a fever*, is a phraseological **but** compositional phrase made up of two main elements:
>
> 1. a semantically autonomous element – *fever* – called *base* of the collocation;
> 2. a bound element – *to run* – called *collocate* of the base; the collocate is said to be bound, or not "free", because its selection by the Speaker in order to express a given meaning depends on the prior selection of the base.

As collocations are modeled in lexical systems by means of standard syntagmatic lexical functions, we will start with a brief presentation of the notion of lexical functions (3.2). We will then proceed with the interpretation of syntagmatic lexical functions as a special type of grammar rules (3.3). Finally (3.4), we will show how such rules can be used to derive a considerable amount syntactic structures embedded in natural language lexicons.

3.2 Standard Lexical Functions

A given standard lexical function is a generalization of a lexical link that possesses the following properties:

- it is either paradigmatic (synonyms, antonyms, nominalizations, verbalizations, actant names, etc.) or syntagmatic (collocates that are intensifiers [***driving*** *rain*], light verbs [***to run*** *a fever*], etc.);
- it is recurrent and universally present in natural languages;
- it is often (though not necessarily) expressed by morphological means (*drive* $\rightarrow$ ***driver*** [actant name], *store* $\rightarrow$ ***mega****store* [intensifier], etc.).

For instance, **Magn** is the standard lexical function that denotes collocational intensifiers; it can be applied to any full lexical unit in order to return the set of all typical intensifiers for that unit.[7] This is illustrated in (1), with the two semantically related units FEVER and HEADACHE as arguments of **Magn**.

(1) a. **Magn**(*fever*) = *high* < *raging*
b. **Magn**(*headache*) = *bad*, *severe* < *terrible*, *violent* < *pounding*, *splitting*

Note that collocative meanings can sometimes be expressed synthetically (within a paradigmatically related term) rather than analytically (as collocates). This phenomenon is call *fusion* and fused values of syntagmatic lexical functions are flagged with the "//" symbol in lexicographic descriptions; for instance:

(2) **Magn**($rain_{\mathrm{V}}$) = *hard*, *heavily*, *//pour down*

Years of lexical studies on a wide spectrum of natural languages have allowed for the identification of a now stable set of approximately 65 *simple lexical functions*;[8] additionally, these functions can be combined to form *complex lexical functions* (Kahane and Polguère, 2001).

The system of lexical functions is a descriptive tool that allows for a rationalization and formalization of the web of paradigmatic and syntagmatic links that connect lexical units in natural languages. This explains why we have adopted lexical functions as the main structuring principle for lexical systems.

3.3 Standard Syntagmatic Lexical Functions as Grammar Rules

We will now focus on standard syntagmatic lexical functions in order to examine how they offer an original treatment of the syntax of collocations. For this, we will use as illustration one specific standard syntagmatic lexical function: **Real$_1$**. It is commonly characterized as follows.

> The lexical function application **Real$_1$**(L) stands for a full verb:
>
> - that expresses such meanings as 'to realize L', 'to do what is supposed to be done as regards to L' ...;

[7] A lexical function is thus quite similar to an algebraic function **f**, that can be applied to a given number x in order to return a given value y: $\mathbf{f}(x) = y$.

[8] The exact number of lexical functions varies according to the descriptive granularity one wants to adopt.

- that takes L as second deep-syntactic actant (i.e. first complement) and the first deep-syntactic actant of L as its first deep-syntactic actant (i.e. grammatical subject).[9]

In case of fusion, the meaning 'L' is encapsulated in the meaning of the lexical function application, together with the sense of realization, and therefore **//Real$_1$**(L) doesn't take L as second syntactic actant.

As an illustration, Figure 5 gives the so-called *article-view* of **Real$_1$** values for BALLOON$_N$ **2** [*We could get there by balloon.*] in the en-LN.[10]

Figure 5: **Real$_1$**(*balloon*$_N$ **2**) in the en-LN.

Standard lexical functions such as **Real$_1$** can be conceptualized from at least two perspectives.

- From the viewpoint of the structure of lexical knowledge, they are universal relations that paradigmatically and syntagmatically connect lexical units within lexical systems.

- From the viewpoint of the universal system of deep-syntactic paraphrasing (Mel'čuk, 2013, Chap. 9), they are "meta lexical units" whose application to a given lexical unit (argument of the lexical function) stands for a set possible lexicalizations in a deep-syntactic structure.

In this latter case, it is important to note that each standard syntagmatic lexical function actually denotes two dependency structures: one for "normal" values of the lexical function application and one for "fused" values. Therefore, the two deep-syntactic trees[11] in Figure 6 are inherently associated to **Real$_1$**.

If we refer to what was said earlier about the lexicon-grammar dichotomy (section 1.1), we are

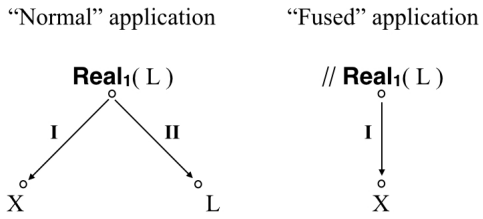

Figure 6: **Real$_1$**'s Deep-syntactic structures.

entitled to consider that trees in Figure 6, because they correspond to general (in this case, universal) linguistic rules about syntactic structuring, are in essence grammatical: they designate syntactic potential that can be run on any lexical rules of the type illustrated in Figure 5 in order to participate in the generation of actual surface-syntactic structures.

3.4 Deriving surface-syntactic structures

In this particular case, rules in Figures 5 and 6 allow for the generation of the three surface-syntactic structures in Figure 7.

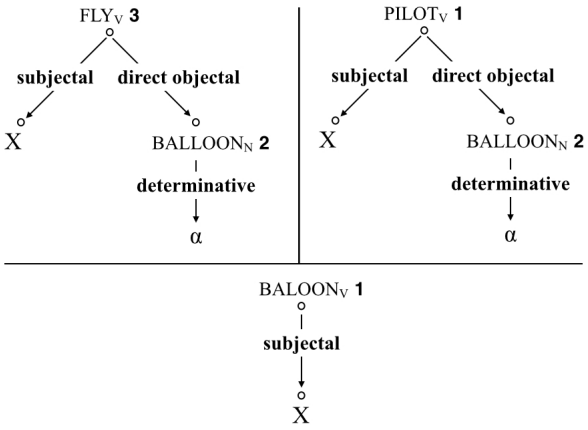

Figure 7: Derived surface-syntactic structures.

If we consider the prospect of such derivation throughout a full lexical system for a given language, we see that a considerable amount of lexicon embedded syntactic structures are extractable from these models. At present, a total number of 7,739 surface-syntactic micro-structures of the type given in Figure 7 can be extracted from the fr-LN.[12] This is of course only a small portion of what is available in the actual French lexicon.

[9] On the notions of semantic and deep-/surface-syntactic actants, see Mel'čuk (2015, Chap. 12).

[10] An article-view, in the lexicographic editor used for building the en- and fr-LNs, is a textual rendering of lexical data associated with a given headword. For details on how lexical function applications are computationally encoded in the en- and fr-LNs, see Gader et al. (2012).

[11] For a concise presentation of Meaning-Text levels of sentence representation and the deep- *vs.* surface-syntax dichotomy, see Kahane (2003).

[12] This corresponds to the number of syntagmatic lexical function relations already woven in the fr-LN.

4 Conclusion: Lexicalized Grammars the Other Way Round

By presenting the syntax of idioms and collocations, we hope to have shown that syntactic information embedded in natural language lexicons goes far beyond phenomena associated to active valency (subcategorization frames). Lexicon embedded syntax is conceptually **and quantitatively** an essential element of lexical knowledge.

It was also our goal to demonstrate that lexical systems such as the fr-LN are particularly suited to the modeling of embedded syntax. In our view, one very promising exploitation of such models for Natural Language Processing (NLP) is the use of large collections of extracted syntactic structures by NLP parsers, for such tasks as disambiguation or processing of phraseological expressions found in corpora.

Collections of syntactic structures extractable from lexical systems bear some conceptual resemblance with *lexicalized grammars* (Schabes et al., 1988), except for the fact that the perspective is totally inverted: rather than lexicalizing grammars, we propose to extract from lexical systems everything actual grammars do not know about syntax.

Acknowledgments

Lexicographic work on the French Lexical Network (fr-LN) originally started at the ATILF CNRS laboratory (Nancy, France) in the context of the RELIEF project funded by the Agence de Mobilisation Économique de Lorraine (AMEL) and the European Regional Development Fund (ERDF).

References

Joseph D. Becker. 1975. The Phrasal Lexicon. In: *Proceedings of the 1975 Workshop on Theoretical Issues in Natural Language Processing (TINLAP'75)*. Association for Computational Linguistics, Cambridge, Mass., 60–63.

Yannick Chudy, Yann Desalle, Benoıt Gaillard, Bruno Gaume, Pierre Magistry and Emmanuel Navarro. 2013. Tmuse: Lexical Network Exploration. In: *The Companion Volume of the Proceedings of IJCNLP 2013: System Demonstrations*. Asian Federation of NLP, Nagoya, 41–44.

Charles J. Fillmore, Christopher R. Johnson and Miriam R. L. Petruck. 2003. Background to FrameNet. *International Journal of Lexicography*, 16(3):235–250.

Nabil Gader, Veronika Lux-Pogodalla and Alain Polguère. 2012. Hand-Crafting a Lexical Network With a Knowledge-Based Graph Editor. In: *Proceedings of the Third Workshop on Cognitive Aspects of the Lexicon (CogALex III)*. The COLING 2012 Organizing Committee, Mumbai, 109–125.

Nabil Gader, Sandrine Ollinger and Alain Polguère. 2014. One Lexicon, Two Structures: So What Gives? In Heili Orav, Christiane Fellbaum and Piek Vossen (eds.): *Proceedings of the Seventh Global Wordnet Conference (GWC2014)*. Global WordNet Association, Tartu, 163–171.

Franz Josef Hausmann. 1979. Un dictionnaire des collocations est-il possible? *Travaux de littérature et de linguistique de l'Université de Strasbourg*, XVII(1):187–195.

Otto Jespersen. 1924. *The Philosophy of Grammar*. George Allen & Unwin, London.

Sylvain Kahane. 2003. The Meaning-Text Theory. In Vilmos Ágel, Ludwig M. Eichinger, Hans Werner Eroms, Peter Hellwig, Hans Jürgen Heringer and Henning Lobin (eds.): *Dependency and Valency. An International Handbook of Contemporary Research*. Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science, de Gruyter, Berlin & New York, 546–569.

Sylvain Kahane and Alain Polguère. 2001. Formal Foundation of Lexical Functions. In: *Proceedings of "COLLOCATION: Computational Extraction, Analysis and Exploitation"*. 39th Annual Meeting and 10th Conference of the European Chapter of the Association for Computational Linguistics, Toulouse, 8–15.

Evelien Keizer. 2007. The lexical-grammatical dichotomy in Functional Discourse Grammar. *Alfa – Revista de Lingüística*, 51(2):35–56.

Veronika Lux-Pogodalla and Alain Polguère. 2011. Construction of a French Lexical Network: Methodological Issues. In: *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop*. Ljubljana, 54–61.

Igor Mel'čuk. 1996. Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon. In Leo Wanner (ed.): *Lexical Functions in Lexicography and Natural Language Processing*. Studies in Language Companion Series 31, John Benjamins, Amsterdam/Philadelphia, 37–102.

Igor Mel'čuk. 2006. Explanatory Combinatorial Dictionary. In Giandomenico Sica (ed.): *Open Problems in Linguistics and Lexicography*. Polimetrica, Monza, 225–355.

Igor Mel'čuk. 2013. *Semantics: From meaning to text*, volume 2. Studies in Language Companion Series 135, John Benjamins, Amsterdam/Philadelphia.

Igor Mel'čuk. 2015. *Semantics: From meaning to text*, volume 3. Studies in Language Companion Series 168, John Benjamins, Amsterdam/Philadelphia.

Igor Mel'čuk, André Clas and Alain Polguère. 1995. *Introduction à la lexicologie explicative et combinatoire*. Duculot, Paris/Louvain-la-Neuve.

Piet Mertens. 2010. Restrictions de sélection et réalisations syntagmatiques dans DICOVALENCE. Conversion vers un format utilisable en TAL. In: *Proceedings of TALN 2010*. Montréal.

Marie-Sophie Pausé. To appear. Modélisation de la structure lexico-syntaxique des locutions au sein d'un réseau lexical. In Maurice Kauffer (ed.): *Actes du colloque international "Approches théoriques et empiriques en phraséologie"*. Eurogermanistik Series, Stauffenburg Verlag, Tübingen.

Alain Polguère. 2009. Lexical systems: graph models of natural language lexicons. *Language Resources and Evaluation*, 43(1):41–55.

Alain Polguère. 2011. Perspective épistémologique sur l'approche linguistique Sens-Texte. *Mémoires de la Société de Linguistique de Paris*, XX:79–114.

Alain Polguère. 2014. From Writing Dictionaries to Weaving Lexical Networks. *International Journal of Lexicography*, 27(4):396–418.

Yves Schabes, Anne Abeillé and Aravind K. Joshi. 1988. Parsing Strategies with 'Lexicalized' Grammars: Application to Tree Adjoining Grammars. In: *Proceedings of the 12th Conference on Computational Linguistics – Volume 2 (COLING'88)*. Association for Computational Linguistics, Budapest, 578–583.