Towards a theoretically-motivated general public dictionary of semantic derivations and collocations for French

Alain POLGUÈRE, Montreal, Canada

Abstract

We present two twin-projects aimed at the description of French semantic derivations and collocations: the DiCo, a formal database whose encoding methodology borrows heavily from explanatory combinatorial lexicology (postulated by Meaning-Text theory), and the *Lexique actif du français* (LAF, 'Active lexicon of French'), which accounts for the same phenomena but is tailored for use by the general public. The LAF is generated entirely from the content of the DiCo, thereby ensuring a rigorous basis to its "popularized" descriptions. Section 1 introduces the general goals of our projects and defines the concepts of semantic derivation and collocation. Section 2 is an example-based presentation of the DiCo; this is mirrored in Section 3, for the LAF. We conclude with remarks on our methodology and information on the progress of the work.

1 Introduction

1.1 The DiCo and the *Lexique actif du français* twin-projects

Our work aims at applying descriptive principles of *EXPLANATORY COMBINATORIAL LEXICOLOGY* (see [MEL'ČUK ET AL. 1995] for a detailed presentation and [MEL'ČUK/POLGUÈRE 1987] for a short introduction) to both the building of lexical databases for natural language processing systems and the writing of general public dictionaries. While the former application has been studied in the past (for instance, in [FONTENELLE 1997, HEID 1996 and IORDANSKAJA ET AL. 1996]), practically nothing has been done on the latter. We will introduce two projects: the *DICO* project targets a computerized lexical database for French, while the *LEXIQUE ACTIF DU FRANÇAIS* (LAF), 'Active lexicon of French') project targets a general public dictionary entirely generated from the DiCo. The research is being conducted jointly by Igor Mel'čuk and the present author, with the help of students from the *Observatoire de linguistique Sens-Texte* (OLST), at the University of Montreal. As both the DiCo and LAF describe mainly semantic derivations and collocations, we need first to clarify these two notions.

1.2 The notions of semantic derivation and collocation

Following Meaning-Text theory's conceptualization of lexical relations, we consider that a relation of *SEMANTIC DERIVATION* holds between two lexical units L_1 and L_2 in any of the following three cases:

1. L_1 and L_2 convey (roughly) the same meaning—L_2 is a (quasi-)synonym, a generic or a conversive of L_1 if it belongs to the same part of speech; otherwise, it is a verbal, nominal, adjectival or adverbial correlate of L_1.

2. L_1 and L_2 have opposite meanings—L_2 is a (quasi-)antonym of L_1.

517

3. L_2 designates an element of the situation designated by L_1—see, e.g., relations between $MURDER_{N/V}$ and the noun MURDERER (standard name of the first actant) or VICTIM (standard name of the second actant).

Such relations between lexical units are called *semantic* derivations since no morphological link needs to exist between the units involved (cf. $MURDER_{N/V}$ and VICTIM), contrary to standard (i.e., morphological) derivation.

Explanatory combinatorial lexicology has identified a small set of semantic derivations found in all natural languages. They are modeled by means of formal tools called *LEXICAL FUNCTIONS*. For instance, the synonymy semantic derivation will be accounted for by means of the **Syn** lexical function which, once applied to the lexical unit L_1—called its *ARGUMENT*, returns the set of units L_2, L_2', L_2'', ...—called the *VALUES* of the application of the lexical function—which are (quasi-)synonyms of L_1. Examples of other lexical functions for semantic derivations are:

- **Gener** for generic terms,
- **Conv** for conversives,
- $\mathbf{S_0}/\mathbf{V_0}/\mathbf{A_0}/\mathbf{Adv_0}$ for nominal/verbal/adjectival/adverbial correlates,
- **Anti** for antonyms,
- $\mathbf{S_1}/\mathbf{S_2}/\ldots$ for names of first/second/... actants.

A given lexical function relation **LF** holding between a lexical unit L_1 and a set of values L_2, L_2', L_2'', ... is formally represented by means of the standard mathematical notation for function application $\mathbf{LF}(L_1) = L_2, L_2', L_2'', \ldots$; e.g., $\mathbf{S_1}(murder) = murderer, killer$. For a detailed presentation of lexical functions and a computational perspective, see [WANNER 1996].

The notion of *COLLOCATION* refers to semi-idiomatic expressions L_1+L_2 such that one of the components, the *COLLOCATE*, is chosen to express a given meaning, in a specific syntactic role, contingent upon the choice of the other component, called the *BASE* of the collocation. As for semantic derivations, collocations are described by means of lexical functions. For instance, intensifiers (*HEAVY bombardment*, *SHARP contrast*, ...), which correspond to the lexical function **Magn**, or support verbs ([to] *RUN a fever*, [to] *STAGE a coup*, ...), which correspond to the lexical functions $\mathbf{Oper_i}$, $\mathbf{Func_i}$ or $\mathbf{Labor_{ij}}$.

It is common practice to say that semantic derivations are encoded by means of *PARADIGMATIC* lexical functions and collocations by means of *SYNTAGMATIC* lexical functions. However, the border between these two families of lexical functions, and hence between the corresponding two types of phenomena, is far from being tight. We will briefly examine two cases illustrating this point.

Firstly, the so-called *FUSED VALUES* of syntagmatic lexical functions—flagged by the "//" symbol in our lexical descriptions—can correspond to semantic derivations. For instance, $FEAST_N$, in one of its senses, is both a fused **Magn** (i.e., intensifier) for MEAL—**Magn**(*meal*) = *big, huge //feast*—and its $\mathbf{Syn_{\supset}}$ (a more specific synonym, i.e., an hyponym)—$\mathbf{Syn_{\supset}}$(*meal*) = *feast*.

Secondly, a $\mathbf{V_0}$ (verbal derivation) for a noun L is linguistically equivalent to the construction $\mathbf{Oper_1}(L)+L$. In other words, the expression $\mathbf{Oper_1}(L)+L$ functions as a paraphrase for $\mathbf{V_0}(L)$. Compare, for example, the two following paraphrases:

X murdered [= $\mathbf{V}_0(murder_N)$] *Y*
X committed [= $\mathbf{Oper}_1(murder_N)$] *Y's murder*

Stretching a bit the concept of fusion, one could say that a $\mathbf{V}_0$ of a lexical unit is at the same time some sort of "fused $\mathbf{Oper}_1$" of this unit.

These two cases demonstrate how intertwined the two phenomena of semantic derivation and collocation are, and why they are considered simultaneously in the DiCo and LAF lexicographic projects, which we will now proceed to describe.

2 The DiCo lexicographic database

2.1 General characterization of the DiCo

The initial aim of the DiCo project was to build a formal lexical database that would meet specific criteria. In terms of content, the database should put the emphasis on the description of combinatorial properties of lexical units, leaving aside lexicographic definitions. In other words, the database should be first and foremost combinatorial rather than explanatory. However, the lexicographic work itself should be based on descriptive principles of explanatory combinatorial lexicology. Because the DiCo describes combinatorial properties of lexical units, its nomenclature should be limited to units which are problematic with regard to such properties. In regard to its form, the DiCo database should be tailored to computer processing: first, it should be automatically translatable into lexicons for natural language processing systems; second, it should be stored in a format that makes it as platform-independent as possible. In short, the DiCo database is to be some sort of a "simplified" and more formalized *EXPLANATORY COMBINATORIAL DICTIONARY* (ECD, see [MEL'ČUK ET AL. 1984, 1988, 1992, 1999]). Because it aims at producing a lexical database that is freely and fully accessible for natural language processing R&D, the DiCo project can be loosely related to the WordNet English database—see [MILLER ET AL. 1990] and [FELLBAUM 1997]. However, there is only a very small intersection between the two projects in terms of the type of data that is accounted for (mainly, synonymy, hyperonymy and meronymy relations, for WordNet) and in terms of encoding (lexical functions, for the DiCo).

The DiCo was initially developed using a home-made dictionary editor; but, as the database grew bigger, it was moved to FileMaker format, without changing much in the actual content of lexical records. It is this FileMaker version of the DiCo that we present here, using a sample entry. Due to lack of space, the sample will have to be a vocable described in the DiCo as monosemic: only one dictionary article/database record has to be considered for this vocable. The vocable in question is Fr. MEURTRE '[a] murder'.

2.2 A sample DiCo record

Each DiCo record describing a lexical unit is structured as a series of eight main fields: 1. name of the unit, 2. grammatical properties, 3. semantic formula, 4. government pattern, 5. synonyms, 6. semantic derivations and collocations, 7. examples and 8. full idioms that include the unit. We list below in text format (Courier font) the content of each of these fields for the MEURTRE

record, adding comments when required. We do not present fields that are used for database management purposes (date/time of modification, lexicographer's ID, etc.).

1. Name of the lexical unit:

```
MEURTRE
```

2. Grammatical properties:

```
nom, masc
```

3. Semantic formula:

```
action de tuer: ~ PAR L'individu X DE L'individu Y
```

The semantic formula is a substitute for an actual lexicographic definition. It starts with a *SEMANTIC LABEL*, as described in [MILIĆEVIĆ 1997], followed by the actantial structure of the lexical unit. The semantic formula above tells us that the core meaning of MEURTRE is 'action de tuer' (an act of killing) and that it involves two actants: the one who kills (X) and the one who is killed (Y). Elements of the formula (other than actant names) that are written in upper case are for human eyes only and do not belong to "pure" DiCo formalization. From a strictly natural language processing point of view, the above formula is to be translated into a structure like

```
action_de_tuer(individu:X, individu:Y).
```

Other components of the formula, such as PAR L' ('by the') and DE L' ('of the'), will prove useful for the generation of the LAF (see Section 3 below).

4. Government pattern:

```
X = I = de N, A-poss
Y = II = de N, A-poss
```

The government pattern indicates (in linear format) the active valency of the lexical unit. This information is presented by means of tables in the published ECDs. The above data tells us that both the killer and the victim of a murder can be expressed as either *de*-prepositional complements or possessive adjectives (French terminology for possessive determiner pronouns). Notice that the DiCo does not mention incompatibilities such as **meurtre de X de Y*.

5. (Quasi-)synonyms:

```
{QSyn} assassinat, homicide#1; crime
```

Synonymy relations have to do with the general organization, structuring and content of the entire database. Therefore, synonyms and quasi-synonyms, including generic terms, are described in a separate field in order to allow us to gather sets of synonyms by activating a simple search procedure. In the above formula, homicide#1 refers to the first sense of the vocable HOMICIDE. The semi-column indicates a relatively important semantic gap between values on the left and the value on the right. In this field, as well as in the next one, lexical relations are encoded by means of lexical functions, whose name appears between curly brackets, before the

list of values; in other words, lexical function applications always have the following format in the DiCo: {<LF>} <Value(s)>.

6. Semantic derivations and collocations:

```
{V0} tuer
{A0} meurtrier-adj
    /*Nom pour X*/
{S1} auteur [de ART ~] //meurtrier-n
    /*Nom pour Y*/
{S2} victime [de ART ~]
    /*Très choquant*/
{Magn} atroce, affreux, brutal, horrible, inqualifiable, odieux
    /*Qui a été préparé*/
{A2Prepar1} avec préméditation, prémédité | postpos //assassinat
{Tel qu'il y a
 deux/trois/quatre Y} double/triple/quadruple | antépos ["Les
                            victimes de ce double meurtre sont un père
                            de famille et son fils de 15 ans."]
    /*Faire un M.*/
{Oper1} accomplir, commettre, perpétrer [ART ~]; tremper [dans
    ART ~] ["Il a refusé de tremper dans ce meurtre odieux."]
    /*Causer que X fasse un M.*/
{CausOper1} pousser [N=X au ~]
    /*Raison d'un M.*/
{S1CausOper1} mobile [de ART ~]
    /*S'occuper d'un M.*/
{Real-I} enquêter [sur ART ~]
{Real-II} élucider [ART ~], trouver l'auteur de ART ~
{Real-III} punir, châtier [ART ~]; venger [ART ~]
    /*Service de police qui s'occupe des M.*/
{S1Real-I/II} brigade criminelle
    /*Préparer un M.*/
{Prepar1} préméditer, préparer [ART ~] //comploter
{Essayer de faire croire
 qu'un M. est un N} maquiller [ART ~ en N] ["Il a maquillé
                            ce meurtre en accident/suicide."]
{Fait de tenter un M.} tentative [de ~]
{Cri lancé par qqn. qui assiste à
 <risque d'être victime d'>un M.} « Au ~ ! »
```

Semantic derivations and collocations are ordered according to standard ECD methodology: paradigmatic lexical functions (which correspond to semantic derivations), followed by syntagmatic lexical functions (which encode collocations). Notice the presence, between /*...*/, of a popularization for most lexical function names, that will be used for the generation of the LAF. For instance, Nom pour X is a popularization for the lexical function S1. As we will see in Section 3, LAF entries do not contain lexical function names; rather, lexical relations

are encoded by means of formulas written in some sort of "meta-French", that is easily understandable even for learners of the language. Some translations of lexical function names into meta-French are rather straightforward (the "☞" symbol for `Syn`, `Nom` for `S0`, `Verbe` for `V0`, `Générique` for `Gener`, etc.) and do not need to be specified; others vary from one entry to another and have to be included by hand in the DiCo record—e.g., here, `Faire un M.` ("to do" a murder), for `Oper1`. We cannot enter into a detailed analysis of all formalisms that are used above, but we believe that most of them are transparent enough to be roughly interpreted without further explanation, specially if the reader has some acquaintance with explanatory combinatorial lexicology.

7. Examples:

```
C'est ici que le double meurtre a été commis. Soupçonné
du meurtre de son épouse, il a été arrêté par les gen-
darmes mercredi. Il devrait comparaître aux assises
dans trois semaines comme auteur présumé du meurtre
d'un quinquagénaire.

------------------------------------------

La mésentente pourrait être le mobile du meurtre.
```

Lexicographic examples are borrowed from TEXTUM, a set of corpora that is used for lexicographic research at the University of Montreal. Eventually, some examples can be slightly edited. (We introduce our own examples in exceptional cases only.) The horizontal line splitting the example field indicates that the last example will not appear in the LAF (where constraints of space apply).

8. Full idioms:

```
_appel au meurtre_
_crier au meurtre_
```

This last field gives pointers to full idioms that formally contain the lexical unit under description and that are themselves actual lexical units, which should have their own record in an expanded DiCo for the whole language. Full idioms are always written in the DiCo between two underscore characters.

This concludes the presentation of a sample DiCo record. The DiCo project has many more important aspects to discuss, but space constraints dictate that we now move to the presentation of the LAF, which is the central topic of this paper.

3 *Lexique actif du français*: Popularizing the DiCo's descriptions

While the concept of lexical function has already been used in natural language processing and computational lexicography, it has never, to our knowledge, materialized in a direct fashion in what can be termed *general public lexicography*. The LAF is a first attempt to bridge the gap

between "theoretical" and "commercial" lexicography with regard to explanatory combinatorial lexicology and, more specifically, to the type of phenomena accounted for by lexical functions.

The objective of the LAF project is to design and compile a general public dictionary of semantic derivations and collocations that would meet the following criteria. In terms of content, it will account for semantic derivations and collocations using Meaning-Text lexicographic principles in a straightforward fashion, thus ensuring an underlying logical structure and theoretical soundness to the description. In terms of form, it will be minimally formal and will contain descriptions that will be as much as possible accessible to a public of non-specialists. As it would be wasteful to multiply lexicographic projects targeting the modeling of identical phenomena, we worked in such a way as to make the generation of the LAF follow exclusively from information contained in the DiCo.

The unique feature of the LAF when compared to equivalent commercial dictionaries—e.g., for French, [LACROIX 1947] and [ROUAIX 1997]—is that it will offer an explicit detailed description of semantic derivations and collocations, rather than just simple listings of lexical links. We believe that a dictionary such as the LAF, even though limited in its nomenclature, would be an invaluable tool for learners of French.

If popularizing scientific concepts is not an easy task, popularizing the product of a scientific description turned out to be an even more challenging one. Unfortunately, we cannot enter here into details of how and why we made the choices we made in determining the format for the LAF. We can however illustrate the LAF with a brief presentation of an entry, using the entry for Fr. MEURTRE below. We encourage the reader to compare this dictionary article with the corresponding DiCo record, analyzed in Section 2.

MEURTRE, nom, masc

> ACTION DE TUER: Meurtre par l'individu X [*de* N, A$_{poss}$] de l'individu Y [*de* N, A$_{poss}$] ☞ assassinat, homicide**1**; crime **VERBE** tuer **ADJECTIF** meurtrier$_{Adj}$ **NOM POUR X** auteur [de ART ~] //meurtrier$_N$ **NOM POUR Y** victime [de ART ~] **TRÈS CHOQUANT** atroce, affreux, brutal, horrible, inqualifiable, odieux **QUI A ÈTÈ PRÉPARÉ** avec préméditation, prémédité | postpos //assassinat **TEL QU'IL Y A DEUX/TROIS/QUATRE Y** double/triple/quadruple | antépos [*Les victimes de ce double meurtre sont un père de famille et son fils de 15 ans.*] **FAIRE UN M.** accomplir, commettre, perpétrer [ART ~]; tremper [dans ART ~] [*Il a refusé de tremper dans ce meurtre odieux.*] **CAUSER QUE X FASSE UN M.** pousser [N$_X$ au ~] **RAISON D'UN M.** mobile [de ART ~] **S'OCCUPER D'UN M.** enquêter [sur ART ~]; élucider [ART ~], trouver l'auteur [de ART ~]; punir, châtier [ART ~]; venger [ART ~] **SERVICE DE POLICE QUI S'OCCUPE DES M.** brigade criminelle **PRÉPARER UN M.** préméditer, préparer [ART ~] //comploter **ESSAYER DE FAIRE CROIRE QU'UN M. EST UN N** maquiller [ART ~ en N] [*Il a maquillé ce meurtre en accident/suicide.*] **FAIT DE TENTER UN M.** tentative [de ~] **CRI LANCÉ PAR QQN. QUI ASSISTE À** <**RISQUE D'ÊTRE VICTIME D'**>**UN M.** « Au ~ ! ». *C'est ici que le double meurtre a été commis. Soupçonné du meurtre de son épouse, il a été arrêté par les gendarmes mercredi. Il devrait comparaître aux assises dans trois semaines comme auteur présumé du meurtre d'un quinquagénaire.* ◇ ⌐appel au meurtre⌐, ⌐crier au meurtre⌐.

Even a quick glance at the above entry suffices to show that, with the LAF, we have taken off from the realm of "theoretical" dictionaries (such as published ECDs) and formal databases (such as the DiCo). This entry, like entries of most commercial dictionaries, can be read at two levels: either in a totally naïve way (without first getting acquainted with the precise meaning of formats and codes), or in a well-informed way—after careful reading of the introduction to

the dictionary. The order of presentation of information follows quite closely what is found in the DiCo. A noticeable exception is the government pattern, which is now dispatched within the actantial description of the lexical unit.

Regarding this part of the entry, one may think that the mention of semantic actants (here, X and Y) is unnecessarily technical and may scare away or confuse the user of the LAF. We do not believe so. First, we found that it was actually impossible to account for the "meaning" of many lexical relations (whether semantic derivations or collocations) in a clear and compact fashion without referring to semantic actants, and actants were better named using Xs and Ys than by saturating our descriptions with expressions such as *the one*, *someone*, *something*, etc. Second, one of the main things the concept of lexical function teaches us is precisely that the phenomena under description are better understood and handled (perhaps, can only be understood and handled) through the mediation of the concept of semantic actant. For instance, once it is stated explicitly that Fr. MEURTRE has two actants, the description of many of its lexical function relations follows naturally:

1. there are values for the standard **NOM POUR X** ('name for X') which are auteur [de ART ~] //meurtrier$_N$ ('author of a murder', 'murderer'), cf. lexical function S_1;

2. there is a standard **NOM POUR Y** ('name for Y') which is victime [de ART ~] ('victim of a murder'), cf. lexical function S_2;

3. there is an idiomatic way to say **CAUSER QUE X FASSE UN M.** ('to cause X to do a murder') which is pousser [N$_X$ au ~] (lit. *to push X to (commit a) murder*), cf. lexical function **CausOper$_1$**.

One big step in popularizing the description was to remove lexical function names and replace them with very general paraphrases in meta-French, as introduced in Section 2. From a scientific point of view, this was a sacrifice because such descriptions can never be as rigorous and falsifiable as formal descriptions. Nevertheless, this sacrifice proved both necessary in the context of a general public dictionary and, most of all, feasible. It appeared that it was not so difficult to paraphrase lexical function formulas with simple, relatively non-idiomatic, and easy to understand expressions in French. Moreover, patterns of translation started to emerge, showing a correlation between expressions in meta-French associated with lexical function formulas and the semantic labelling we had performed on lexical units. We will return to this interesting question in Section 4 below. In the case of semantic actants however, for reasons as highlighted above, we do not believe that at the end of the day the description would gain in clarity if explicit references to them were dropped.

4 Remarks on methodology and current state of the projects

Our lexicographic twin-projects have now entered the development phase, with the elaboration of a "production line" by which the same team can quasi-simultaneously build two very different lexicographic products. Starting from a pre-selected nomenclature of about 2,500 vocables (chosen by a thorough examination of all French lexical units that seem to control a significant set of lexical relations), the following tasks are sequentially performed:

1. rough drafting of DiCo entries where main senses are distinguished and examples are gathered from corpora (students' help is invaluable here),

2. entering of semantic formulas plus encoding of semantic derivations and collocations by means of lexical functions,

3. cleaning up of the entries and introduction of popularizing formulas as comments in the DiCo,

4. manual generation of LAF entries from text file versions of the DiCo records (mainly, application of predefined paragraph and character styles performed with FrameMaker),

5. revision of the LAF version of entries, which entails corrections in the LAF itself and in the DiCo.

Our experience is that descriptive problems (missing information, wrongly encoded links, etc.) show better in LAF entries whereas solutions to these problems are better found by examining the DiCo records, which contain lexical function encoding. We are not sure yet what the practical and theoretical implications of this observation are, but we feel that they will turn out to be far from trivial. Among other things, we found that the standard way of listing lexical function relations in the ECD, using lexical function names as keywords, could sometimes give poor results once the encoding is translated into popularized form. For instance, we have the following sequence of lexical function relations in the DiCo record given in Section 2:

```
        /*Très choquant*/
{Magn} atroce, affreux, brutal, horrible, inqualifiable, odieux
        /*Qui a été préparé*/
{A2Prepar1} avec préméditation, prémédité | postpos //assassinat
```

This is to ensure that collocates that are modifiers of the base will all appear together in the LAF entry: *meurtre atroce/affreux/brutal/*...next to *meurtre avec préméditation/prémédité*. However, in the standard ECD approach, these two sets of collocations would have been listed far apart because the second one is conceived of as being related to the **Prepar** lexical function: verb that takes the keyword as first complement and expresses the general meaning 'to prepare'. Because they correspond to verbal collocates, values of **Prepar** are normally listed with other such collocates. For instance, in our sample DiCo record, the following lexical function relation

```
        /*Préparer un M.*/
{Prepar1} préméditer, préparer [ART ~] //comploter
```

appears after relations that belong to the **Oper** and **Real** lexical function families. The DiCo is to be used to derive both lexicons for natural processing systems and general public dictionaries (such as the LAF). The order of presentation of lexical function relations is irrelevant in the first case (where computer programs can access data regardless of their order of presentation in a database). However, it is crucial in the context of a dictionary such as the LAF. Because we want the DiCo to be translated in a systematic fashion into LAF format, we have chosen to adopt in the DiCo an order of presentation that meets the LAF requirements, rather than following strictly ECD standards, that are fit only for users trained in explanatory combinatorial lexicography.

525

A very interesting topic which ought to be discussed in a lengthier presentation is the determination of popularizing formulas for lexical function relations. Our ultimate goal is to give as much as possible of a logical basis to the selection and use of such formulas. A powerful strategy to follow is to ensure a certain homogeneity between these formulas and the semantic label associated with each lexical unit. For instance, it appears that intensifiers (lexical function **Magn**) of lexical units labelled with SENTIMENT ('feeling') can be paraphrased with the formula **INTENSE** ('intense'), whereas **MARQUÉ** ('marked') is a better choice for units labelled with APPARENCE ('appearance'). At this point of our research, we do not yet possess a rigorous and well-defined methodology for selecting popularizing formulas. We are confident however that such a methodology will be one of the outcomes of our work and that it is a first step toward the determination of a meta-French for the encoding, understanding, and maybe teaching of lexical function relations.

We hope to have brought to their term both the DiCo and LAF projects by the end of this year. Our aim is to publish the LAF in hardcopy as well as in electronic format. Significant samples of the DiCo database can be obtained from the author of the present paper in FileMaker Pro or tab-separated text format.

Acknowledgments

References

[FELLBAUM 1997] Fellbaum, C. (ed.) (1997). *WordNet: an electronic lexical database*. MIT Press, Cambridge, Mass.

[FONTENELLE 1997] Fontenelle, T. (1997). *Turning a Bilingual Dictionary into a Lexical-Semantic Database*. Lexicographica Series Maior, Vol. 79, Max Niemeyer, Tübingen.

[HEID 1996] Heid, U. (1996). "Using Lexical Functions for the Extraction of Collocations From Dictionaries and Corpora", in: [WANNER 1996], pp. 115-146.

[IORDANSKAJA/KIM/POLGUÈRE 1996] Iordanskaja, L., Kim, M. and Polguère, A. (1996). "Some Procedural Problems in the Implementation of Lexical Functions for Text Generation", in: [WANNER 1996], pp. 279-297.

[LACROIX 1947] Lacroix, U. (1947). *Dictionnaire des mots et des idées: les idées par les mots*. Nathan, Paris.

[MEL'ČUK ET AL. 1984] Mel'čuk, I. A., Arbatchewsky-Jumarie, N., Elnitsky, L., Iordanskaja, L. and Lessard, A. (1984). *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexico-sémantiques I*. Presses de l'Université de Montréal, Montréal.

[MEL'ČUK ET AL. 1988] Mel'čuk, I. A., Arbatchewsky-Jumarie, N., Dagenais, L., Elnitsky, L., Iordanskaja, L., Lefebvre, M.-N. and Mantha, S. (1988). *Dictionnaire explicatif et combinatoire du*

français contemporain: Recherches lexico-sémantiques II. Presses de l'Université de Montréal, Montréal.

[MEL'ČUK ET AL. 1992] Mel'čuk, I. A., Arbatchewsky-Jumarie, N., Iordanskaja, L. and Mantha, S. (1992). *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexico-sémantiques III*. Presses de l'Université de Montréal, Montréal.

[MEL'ČUK ET AL. 1999] Mel'čuk, I. A., Arbatchewsky-Jumarie, N., Iordanskaja, L., Mantha, S. and Polguère, A. (1999). *Dictionnaire explicatif et combinatoire du français contemporain: Recherches lexico-sémantiques IV*. Presses de l'Université de Montréal, Montréal.

[MEL'ČUK/CLAS/POLGUÈRE 1995] Mel'čuk, I. A., Clas, A. and Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. AUPELF UREF/Duculot, Louvain-la-Neuve.

[MEL'ČUK/POLGUÈRE 1987] Mel'čuk, I. A. and Polguère, A. (1987). "A Formal Lexicon in the Meaning-Text Theory: or How to Do Lexica with Words", in: *Computational Linguistics*, Vol. 13/3-4, pp. 261-275.

[MILIĆEVIĆ 1997] Milićević, J. (1997). *Étiquettes sémantiques dans un dictionnaire formalisé du type Dictionnaire Explicatif et Combinatoire*. M.A. dissertation, Département de linguistique et de traduction, Université de Montréal.

[MILLER ET AL. 1990] Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D. and Miller, K. J. (1990). "Introduction to WordNet: An On-line Lexical Database", in: *International Journal of Lexicography*, Vol. 3/4, pp. 235-244.

[ROUAIX 1997] Rouaix, P. (1997). *Trouver le mot juste Ů Dictionnaire des ides idées suggérées par les mots*. Le Livre de Poche no. 7939, Armand Colin, Paris.

[WANNER 1996] Wanner, L. (ed.) (1996). *Lexical Functions in Lexicography and Natural Language Processing*. Benjamins, Amsterdam/Philadelphia.