

# Critères heuristiques pour l'encodage des collocations au moyen de fonctions lexicales

Margarita ALONSO RAMOS, A Coruña, Spain

## Résumé

Lexical functions (LFs), as developed by I. Mel'čuk in the Meaning-Text Theory are a formal means for the representation of collocations. LFs allow for a fine-grained description of collocations with respect to their syntactic and semantic characteristics. However, it has repeatedly been pointed out that the definition of LFs is far from being transparent for linguists and lexicographers not trained in this theoretical framework. Especially problematic from this perspective is the combinatorial capacity of LFs. In this paper, we suggest some criteria for the encoding of collocations in terms of LFs. These criteria draw on LF definitions, i.e., both the syntactic and semantic properties of LFs. The research reported on in the paper is part of an ongoing project within which an on-line collocational dictionary of Spanish is compiled. Our objective is to incorporate a heuristic model the backbone of which are the criteria for the encoding of LFs into an interface of the lexical database. The interface will support the identification of (combinations of) LFs and permit the verification of proposed formulations.

## 1 Introduction : les fonctions lexicales et leur encodage

Les fonctions lexicales, définies dans le cadre de la Théorie Sens-Texte (TST) (voir [ZHOLKOVSKY/MEL'ČUK 1967, MEL'ČUK 1982, WANNER 1996]), sont déjà connues depuis un certain temps comme un outil puissant pour décrire les relations sémantico-syntaxiques, paradigmatiques aussi bien que syntagmatiques entre les lexies. *Grosso modo*, une fonction lexicale (= FL) est un sens assez abstrait ou un certain rôle sémantico-syntaxique tel que son expression linguistique dépend du lexème auquel cette FL s'applique. Ainsi par exemple, la FL **Magn** ('intense') appliquée au mot-clé *mémoire* rend la valeur *prodigieuse*, tandis qu'appliquée à *désir* elle rend la valeur *ardent*. Il existe environ soixante FL qui décrivent des collocations nominales, adjectivales, verbales et prépositionnelles aussi bien d'un point de vue sémantique que syntaxique.

Or, les FL ne sont pas toujours définies explicitement, particulièrement en ce qui concerne leur capacité combinatoire. Les FL peuvent se combiner entre elles pour former des FL complexes, des FL composées et des configurations des FL (voir [MEL'ČUK 1996] et [ALONSO RAMOS/TUTIN 1996]). Le résultat de ces combinaisons donne souvent une chaîne de FL dont la syntaxe interne est loin d'être transparente pour le lecteur non initié au cadre lexicologique de la TST. Comme ROBIN [1990] l'a signalé : "... although the notion of lexical function in itself is appealing [...], using the very LFs proposed by Mel'čuk is impractical without the collaboration of an MTT guru".

Ainsi, peu de lexicographes débutants dans le formalisme des FL pourraient formuler la combinaison suivante de FL qui apparaît sous l'entrée ATTENTE I.2 [MEL'ČUK ET AL. 1999] :

**A<sub>1</sub>AntiBonInvolv<sub>1</sub> + Oper<sub>1</sub>** : subir [des ~ s] | A. au pl. et a un modificateur

Ex. : *Les banlieusards seront condamnés à subir de longues attentes à cause des travaux.*

La glose en langue naturelle d'une telle combinaison de FL serait approximativement : 'l'attente affectant négativement X, X est en attente'.

Il n'existe pas des critères précis qui empêchent un lexicographe d'encoder une collocation donnée par plus d'une combinaison de FL ; c'est-à-dire que différentes combinaisons de FL peuvent représenter le sens d'une collocation donnée et il n'y a pas, pour le moment, de guides qui permettent de choisir la meilleure combinaison. Ainsi, pour l'exemple précédent, on pourrait avoir noté :

**Conv<sub>21</sub>(AntiBonInvolv<sub>1</sub>)** : subir [des ~ s] | A. au pl. et a un modificateur

Dans ce cas, la FL **Conv**(ersion) permute l'ordre des actants de l'autre combinaison de FL. Ainsi, si **AntiBonInvolv<sub>1</sub>** est glosé par 'l'attente affecte négativement X', le résultat de la conversion nous donne 'X est affecté négativement par l'attente', qui est aussi, *grosso modo*, le sens de *subir des attentes*.

Dans le cadre du *Diccionario de colocaciones y marcadores del español* (DICME), que nous développons actuellement, nous nous proposons d'établir certains critères heuristiques afin d'aider le lexicographe pour l'encodage des collocations par les FL.<sup>1</sup> Ces critères doivent découler d'une étude profonde du sens et de la syntaxe de toutes et de chacune des FL, travail que nous avons entrepris dans [ALONSO RAMOS 1993].

Dans la section suivante, nous ébaucherons de tels critères d'abord pour les FL simples et ensuite pour les combinaisons de FL. Dans la dernière section, nous montrerons comment ces critères peuvent être appliqués pour l'encodage d'une collocation donnée au moyen d'une combinaison de FL.

## 2 Critères heuristiques pour l'encodage de FL

Pour établir des critères qui guident l'encodage de FL, il faut d'abord procéder à une description précise de toutes les FL. Pour ce faire, nous avons employé des paramètres sémantiques et syntaxiques, en spécifiant quelle est la relation sémantique et syntaxique entre le mot-clé et la valeur. Nous avons aussi classifié les FL comme paradigmatiques ou comme syntagmatiques et nous avons noté quelle est la capacité combinatoire de chacune (voir [ALONSO RAMOS 1993] et [ALONSO RAMOS/TUTIN 1996]).

### 2.1 Critères concernant les FL simples

Voici maintenant quelques critères qui doivent servir à l'encodage de FL simples. L'objectif est d'associer à la base de données un outil qui aide à diagnostiquer, par le biais d'une fenêtre de dialogue, la FL en cause. Devant une possible FL, le lexicographe doit vérifier les faits suivants :

1. Le mot-clé et la valeur cooccurrent-ils ou non ? Si oui, la FL est syntagmatique. Si non, la FL est paradigmatique ou c'est une expression fusionnée d'une FL syntagmatique ;

2. Quelle est la relation sémantique entre le mot-clé et la valeur : dérivation sémantique,<sup>2</sup> d'un côté et relation de prédicat-argument, de l'autre ? Si le mot-clé et la valeur sont des dérivés sémantiques, la FL candidate appartiendra aux FL paradigmatiques. Si la relation entre la valeur et le mot-clé est de prédicat-argument, la FL candidate sera une FL syntagmatique ;
3. Quelle est la classe syntaxique du mot-clé et quelle est celle de la valeur ? Si le mot-clé est un nom et la valeur est un adjectif, les FL candidates seront **Magn, Bon, Ver, Epit** et **Pos** ;
4. Quelle est la relation syntaxique entre le mot-clé et la valeur ? Si le mot-clé est un actant syntaxique de la valeur, les FL candidates seront des FL verbales, nominales et prépositionnelles.

Dans la Figure 1, nous proposons une fenêtre d'identification d'une FL simple dans laquelle ces quatre considérations sont prises en compte. Le mot-clé est nommé  $L_1$  et la valeur,  $L_2$ . Les deux premières considérations nous aident faire le choix entre le groupe des FL paradigmatiques et celui des FL syntagmatiques. Dans les deux dernières, nous exemplifions les FL candidates possibles. Par exemple, si le mot-clé  $L_1$  a une relation syntaxique modificative (notée **ATTR**) avec la valeur  $L_2$ , les FL candidates sont **Magn, Ver, Bon, Epit** et **Pos**. Si le mot-clé  $L_1$  est l'actant syntaxique **I/II/III** (noté par un chiffre romain), le choix des FL candidates est restreint.

Examinons quelques exemples. Pour encoder la relation entre *crime* et *criminel*, par exemple, la première chose à faire est de vérifier si les deux lexies cooccurrent. Si ce n'est pas le cas, le choix de la FL candidate pointe vers les FL paradigmatiques.<sup>3</sup> Ensuite, il faut vérifier quelle est la relation sémantique entre les deux lexies. Étant donné que *criminel* désigne un participant dans la situation nommée par *crime*, le choix vise encore les FL paradigmatiques. Notons que même si *criminel* est le nom typique pour le premier actant de *crime* ('celui qui commet le crime'), on ne peut pas dire qu'il y ait une relation de prédicat-argument entre les deux lexies : 'crime' ne prédique rien de 'criminel' et encore moins, 'criminel' ne prédique de 'crime'.<sup>4</sup>

Analysons un autre cas où les deux lexies cooccurrent comme *alimentation équilibrée*. Cette fois, la relation sémantique entre la valeur et le mot-clé est de prédicat-argument : 'équilibrée (alimentation)'. Le choix de la FL candidate pointe vers les FL syntagmatiques. Puisque le mot-clé est un nom et la valeur est un adjectif, le choix est déjà plus restreint : il doit s'agir d'une FL modificatrice comme **Magn, Ver, Bon, Pos** ou **Epit**. Le critère de la relation syntaxique coïncide avec le critère de la classe syntaxique du mot-clé et de la valeur : le mot-clé *alimentation* gouverne syntaxiquement la valeur *équilibrée*, ce qui serait représentée dans la syntaxe de la TST comme :

ALIMENTATION -**ATTR** → ÉQUILIBRÉE<sup>5</sup>

Les seules FL candidates qui peuvent répondre à cette configuration sont les mêmes que celles prévues par le critère de la classe syntaxique.

Il nous reste à montrer une collocation dans laquelle la relation syntaxique entre la valeur et le mot-clé est actancielle. Parmi les collocations N + N, prenons, par exemple, *chef d'orchestre*. D'abord, la cooccurrence et la relation sémantique nous poussent vers les FL syntagmatiques. Un chef est toujours chef de quelque chose, donc on a une relation de prédicat-argument : 'chef(orchestre)'. Le critère de la classe syntaxique nous oriente vers les FL **Sing, Mult, Cap**,

<b>FAITS À CONSIDÉRER</b>	<b>FL CANDIDATES</b>
<b>1. Cooccurrence ou non entre L<sub>1</sub> et L<sub>2</sub></b>	
– Oui .....	FL syntagm : Oper, Func, Labor Real, Fact, Labreal Magn, Bon, Ver, etc. Sing, Mult, Cap, etc.
– Non .....	FL paradigm : Syn, Anti, Conv S <sub>0</sub> , A <sub>0</sub> , V <sub>0</sub> , Adv <sub>0</sub> S <sub>1</sub> , A <sub>1</sub> , Able <sub>1</sub> , etc.
<b>2. Relation sémantique entre L<sub>1</sub> et L<sub>2</sub></b>	
– Dérivation sémantique .....	FL paradigmatiques
– Prédicat-argument .....	FL syntagmatiques
<b>3. Classe syntaxique de L<sub>1</sub> et de L<sub>2</sub></b>	
<b>L<sub>1</sub></b>	<b>L<sub>2</sub></b>
N	N ..... Syn, Sing, Cap, etc. A ..... A <sub>0</sub> , Magn, Ver, etc. V ..... V <sub>0</sub> , Oper, Real, etc. Adv ..... Loc, Adv, Propt, etc.
<b>L<sub>1</sub></b>	<b>L<sub>2</sub></b>
A	N ..... S <sub>0</sub> A ..... Syn, Anti, Conv, etc. V ..... Pred Adv ..... Magn, Bon, Ver, etc.
<b>L<sub>1</sub></b>	<b>L<sub>2</sub></b>
V	N ..... S <sub>0</sub> V ..... Syn, Anti, Conv, etc. Adv ..... Loc, Adv, Propt, etc.
<b>4. Relation syntaxique entre L<sub>1</sub> et L<sub>2</sub></b>	
– <b>Relation modificative</b>	
L <sub>1</sub> -ATTR→L <sub>2</sub> .....	Magn, Ver, Bon, etc.
L <sub>2</sub> -ATTR→L <sub>1</sub> .....	Gener, Figur
– <b>Relations actantielles</b>	
L <sub>2</sub> -I→L <sub>1</sub> .....	Func, Involv, Manif, etc.
L <sub>2</sub> -II→L <sub>1</sub> .....	Oper, Real, Sing, Mult
L <sub>2</sub> -III→L <sub>1</sub> .....	Labor, Labreal

FIG. 1: Fenêtre d'identification d'une FL simple

etc. L'examen de la relation syntaxique entre la valeur et le mot-clé pointe aussi vers les mêmes FL. La valeur *chef* gouverne syntaxiquement le mot-clé, ce qui est représenté ainsi pour la configuration de la FL **Cap** : CHEF -II → ORCHESTRE.

À la suite de cette première vérification, il peut encore rester plusieurs FL candidates mais ces quatre premiers critères auront déjà restreint considérablement les possibilités. En ce qui concerne la classe syntaxique du mot-clé, il existe des règles sans exceptions qui empêchent de proposer, par exemple, une FL comme **V<sub>0</sub>** (verbe dérivé) pour un mot-clé verbal. Les noms acceptent sans aucune exception toutes les FL, alors que toutes les autres classes syntaxiques ont des restrictions. Ainsi, les verbes ne peuvent pas être mot-clé des FL syntagmatiques dont la valeur est un verbe ou une préposition. Également, en ce qui concerne la relation syntaxique entre le mot-clé et la valeur, il n'y a que deux FL simples qui répondent à la configuration syntaxique  $L_2$  -III →  $L_1$ . Seulement **Labor** et **Labreal** peuvent être la tête syntaxique du mot-clé comme troisième actant. Par exemple, la relation syntaxique entre la valeur et le mot-clé de la FL **Labor**<sub>12</sub>(*torture*) = *soumettre* [N à ~] serait représentée ainsi :

SOUMETTRE -III → TORTURE

## 2.2 Critères concernant les combinaisons de FL

Un des attraits des FL, qui cependant les rend aussi très problématiques, est leur grande capacité combinatoire. Il existe plusieurs types de combinaisons de FL. Nous allons nous concentrer seulement sur les relations entre les FL constituantes d'une FL complexe.

Comme pour les relations syntaxiques entre le mot-clé et la valeur, les relations syntaxiques entre deux FL **F** et **G** peuvent être seulement de deux types : modificative (notée par **ATTR**) ou actancielle (noté par un chiffre romain). La FL qui est le gouverneur syntaxique de la relation sera appelée *FL TÊTE* ; la FL qui est le dépendant de la relation **ATTR** sera appelée *FL MODIFICATRICE* et la FL qui est liée par une relation actancielle, sera nommée *FL ACTANCIELLE* (voir Fig. 2).

Quant aux FL liées syntaxiquement par une relation modificative, n'importe quelle FL verbale (à l'exception des verbes supports) peut être la FL tête de la FL complexe. Ainsi, la collocation *se graver dans la mémoire*, comme dans *Ses paroles se gravèrent dans ma mémoire* serait encodée par une FL complexe dont les deux FL constituantes seraient liées par une relation syntaxique modificative :

**MagnReal**<sub>2</sub>(*mémoire*) = *se graver* [dans ART ~]

La FL tête est la FL verbale puisque la valeur est un verbe. La FL modificatrice **Magn** sert à intensifier la FL verbale. La glose en langue naturelle est approximativement : 'l'information Y est très fixée dans la mémoire'.

Par rapport aux FL complexes, liées syntaxiquement par une relation actancielle, la FL tête est restreinte aux FL phasiques et aux FL causatives. Il n'y a pas de restriction, cependant, pour la FL actancielle. Par exemple, la collocation *remettre N en mémoire*, comme dans *Cette musique me remet le passé en mémoire*, sera encodée par une FL complexe dont les deux FL constituantes sont liées par une relation actancielle :

SYNTAXE ENTRE F ET G	FL CANDIDATES
<b>1. Relation modificative entre F et G (G←ATTR-F)</b>	
– F tête .....	FL verbale : – Oper, Func, Labor Real, Fact, Labreal Involv, Manif, Degrad
.....	FL nominale : S <sub>i</sub>
– G modificatrice .....	Fl modificat. : Magn, Bon, Ver
<b>Exemples :</b>	
MagnReal <sub>2</sub> ( <i>mémoire</i> ) = <i>se graver</i> [dans ART ~]	
MagnS <sub>2</sub> ( <i>coûter</i> ) = <i>une fortune</i>	
<b>2. Relation actantielle entre F et G (F-II→G)</b>	
– F tête .....	FL phasique : Incep, Cont, Fin
.....	FL causative : Caus, Liqu, Perm
– G actantielle .....	FL verbale (sans restriction)
<b>Exemples :</b>	
FinOper <sub>1</sub> ( <i>appétit</i> ) = <i>perdre</i> [l'~]	
CausReal <sub>2</sub> ( <i>mémoire</i> ) = [Cl <sub>dat</sub> ] <i>remettre</i> [N en ~]	

FIG. 2: Fenêtre d'identification de la syntaxe d'une FL complexe

**CausReal<sub>2</sub>**(*mémoire*) = [Cl<sub>dat</sub>] *remettre* [N en ~]

La FL tête est **Caus** et son deuxième actant est **Real**. La glose en langue naturelle est approximativement : 'W cause que l'information Y revient en mémoire'.

Comme nous le voyons, si le sens de la collocation à coder inclut le sens 'causation', on choisit une FL de la famille de **Caus** comme FL tête **F** de la FL complexe. La FL **G** sera choisie, d'abord, avec des critères sémantiques, et ensuite, avec des critères syntaxiques. Si l'événement causé est seulement son existence, **G** sera une FL de verbe support. Si c'est la réalisation d'un événement ou l'utilisation d'un artefact, on choisira une FL de la famille de **Real**. S'il s'agit d'une manifestation, on choisit **Manif** et ainsi de suite. Une fois la deuxième FL sélectionnée sémantiquement, on raffine le choix avec des critères syntaxiques. Par exemple, pour choisir entre **Oper** et **Func** comme FL actanciennes, on examine l'ordre des actants de la collocation à coder. Si le mot-clé est le troisième complément de la valeur, **Caus** se combinera avec **Oper**, comme dans **Caus<sub>2</sub>Oper<sub>1</sub>**(*colère*) = *mettre* [N en ~]. Si le mot-clé est le deuxième complément de la valeur, **Caus** se combine avec **Func**, comme dans **Caus<sub>2</sub>Func<sub>1</sub>**(*colère*) = *susciter* [la ~ chez N].

### 3 Mise en pratique des critères d'encodage de FL

Pour montrer comment on peut mettre en pratique les critères que nous venons d'ébaucher, nous suggérons un exercice de formulation de FL. Voici le chemin que devrait suivre le lexicographe cherchant à décrire la collocation *mort de fatigue* par une FL ou une combinaison de FL :

1. Quelle lexie est sélectionnée d'après son sens ? La réponse est *fatigue*, donc c'est la lexie *fatigue* qui est mot-clé de la FL ou la base de la collocation.
2. Quelle est la relation syntaxique entre le mot-clé  $L_1$  et la valeur  $L_2$ , modificative ou actancielle ? La réponse est modificative car *mort* et *fatigue* ne soutiennent aucune relation actancielle.
3. Quelle est la tête syntaxique, le mot-clé  $L_1$  ou la valeur  $L_2$  ? La réponse est que la valeur *mort* est la tête qui gouverne syntaxiquement *de fatigue*. La représentation syntaxique correspondante est donc :

MORT -ATTR→DE-II→FATIGUE

Ce choix élimine le groupe des FL modificatrices (**Magn**, **Bon**, **Ver**) comme seule FL pour décrire notre collocation, car la syntaxe de ces FL exige le mot-clé comme tête syntaxique. La valeur *mort* est un adjectif qui renvoie au premier actant, celui qui éprouve la fatigue. La FL tête doit être adjectivale et paradigmatisée pour inverser la relation syntaxique entre le mot-clé et la valeur.

4. Laquelle des FL modificatrices participe dans la collocation, **Epit**, **Bon**, **Ver**, **Pos** ou **Magn** ?
  - 4.1 **Epit** donne des adjectifs sémantiquement vides. Cependant, *mort* ajoute un sens.
  - 4.2 **Bon**, **Ver** et **Pos** produisent des adjectifs avec une composante d'évaluation. Cependant, dans *mort de fatigue*, il n'y a pas telle composante.
  - 4.3 Le choix est **Magn** car cette FL sert à intensifier le sens du mot-clé. Dans notre collocation, on a une composante sémantique qui intensifie 'fatigue', donc **Magn** doit apparaître dans la formule de FL.
5. Laquelle des FL adjectivales paradigmatisées fait référence à un actant, **Able<sub>i</sub>**, **Qual<sub>i</sub>** ou **A<sub>i</sub>** ?
  - 5.1 Les FL **Able<sub>i</sub>** et **Qual<sub>i</sub>** sont exclues car ces FL produisent des adjectifs avec un sens de probabilité que l'adjectif *mort* n'a pas.
  - 5.2 La seule option valable comme FL tête est **A<sub>i</sub>**, car nous avons affaire à un adjectif qui caractérise un actant de *fatigue*.
6. Quel est l'actant de *fatigue* auquel fait référence la valeur *mort*, le premier **I** ou le deuxième **II** ? Étant donné que dans *mort de fatigue*, nous faisons allusion à celui qui éprouve la fatigue, le seul choix est **A<sub>1</sub>**.  
Il ne reste qu'à combiner les deux FL candidates **Magn** et **A<sub>1</sub>**. Le seul choix possible est une configuration : **Magn + A<sub>1</sub>**. Examinons la justification de ce choix.
7. Comment sont combinées les deux FL choisies, **Magn** et **A<sub>1</sub>** ?

7.1 Examen de la possibilité qu'elles forment une FL complexe : **MagnA<sub>1</sub>**

Pour former une FL complexe, les deux FL doivent être syntaxiquement liées. Si l'on prend **A<sub>1</sub>** comme FL tête, **Magn** devrait être la modificatrice. Cependant, **\*MagnA<sub>1</sub>** n'est pas correcte. Cette combinaison produirait un adjectif qui intensifie l'attribut typique du premier actant ; quelque chose bizarre comme 'très tel qu'il éprouve de la fatigue'. On voit donc que ces deux FL ne sont pas syntaxiquement liées. La langue ne permet pas une relation comme

**\*Magn←ATTR-A<sub>1</sub>**

D'ailleurs, **Magn** ne peut pas être la FL tête car la syntaxe de la collocation n'est pas celle qui décrit la FL **Magn**. Dans notre collocation, la valeur gouverne syntaxiquement le mot-clé, tandis que dans les collocations où **Magn** est la FL tête, c'est le mot-clé qui gouverne la valeur.

7.2 Examen de la possibilité qu'elles forment une FL composée : **A<sub>1</sub>(Magn)**

Pour former une FL composée, une FL doit s'appliquer à la valeur de l'autre FL. Étant donné que la FL **A<sub>1</sub>** est la FL tête, elle devrait être la dernière à être appliquée. La FL composée **\*A<sub>1</sub>(Magn(*fatigue*))** n'est pas non plus correcte. Une telle combinaison devrait être décomposée en **Magn(*fatigue*) = grand, énorme, extrême** et en **A<sub>1</sub> (*grand/énorme/extrême*)**. Évidemment, cette dernière FL ne nous donne pas *mort*. D'ailleurs, cette FL ne serait pas bien formulée car **A<sub>1</sub>** ne prend pas des adjectifs comme mot-clé.

7.3 Examen de la possibilité qu'elles forment une configuration de FL : **Magn + A<sub>1</sub>**.

Pour former une configuration de FL, les deux doivent être appliquées au mot-clé sans qu'il y ait une relation syntaxique entre elles. Étant donnée que la FL **A<sub>1</sub>** est la FL tête, elle doit être notée à la droite de la combinaison. Nous arrivons à la configuration **Magn + A<sub>1</sub>(*fatigue*)**. La FL **Magn** s'applique directement au mot-clé pour l'intensifier et **A<sub>1</sub>** s'applique aussi pour produire un adjectif qui qualifie au premier actant du mot-clé. On a finalement 'tel qu'il éprouve [**A<sub>1</sub>**] une fatigue intense [**Magn**]'.

Le chemin que nous avons suivi pour encoder la collocation *mort de fatigue* par la FL complexe **Magn + A<sub>1</sub>(*fatigue*)** peut être modélisé par un arbre de décision que nous présentons dans la Fig. 3.

Pour conclure, nous voulons mettre l'emphase sur l'idée que la manipulation des FL pour l'encodage des collocations nécessite des connaissances de type "connaissances d'expert". Nous sommes encore loin de pouvoir modéliser ces connaissances par un programme informatique de type système expert, mais nous croyons que les critères ébauchés ici peuvent servir comme premier "mode d'emploi des FL" pour l'encodage des collocations dans notre projet du DICME.

## Remerciements

Cette recherche est menée dans le cadre du projet de recherche intitulé "Base de datos léxico-sémantica : unidades léxicas descriptivas y no descriptivas" du Programme PGIDT99PXI10401B financé par la *Secretaría Xeral de Investigación e Desenvolvemento* (Xunta de Galicia, España).

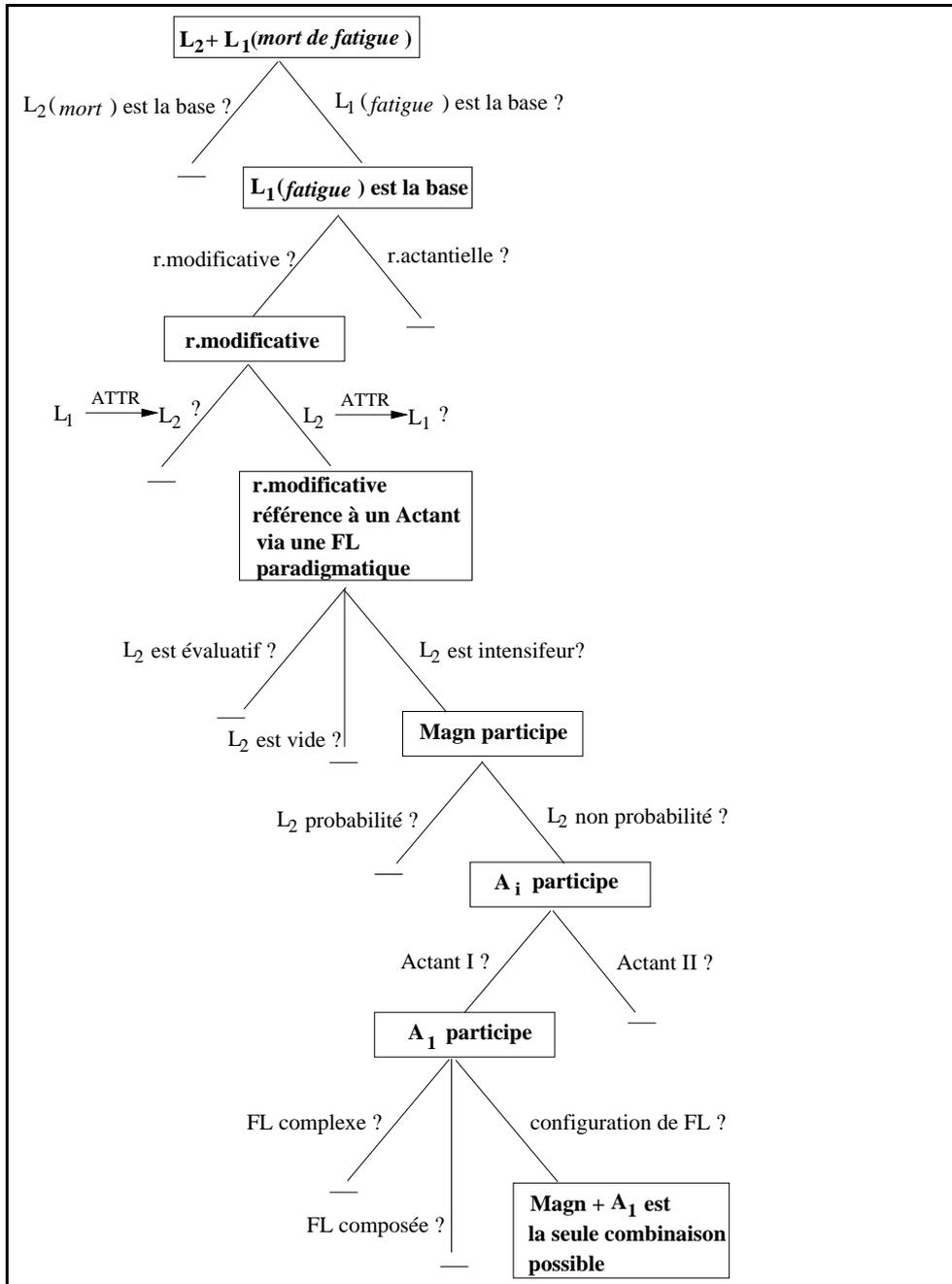


FIG. 3: Arbre de décision pour encoder une FL complexe Magn + A<sub>1</sub>

## Notes

<sup>1</sup>Le DICME est conçu comme une base de données lexicales, comme une lexique informatisée qui met l'emphase sur la cooccurrence lexicale restreinte, spécialement en ce qui concerne les lexies descriptives. Pour la distinction entre les lexies descriptives et non descriptives, voir [IORDANSKAJA/MEL'ČUK 1995]. Le développement des articles pour des lexies non descriptives (connecteurs, marqueurs discursifs, interjections etc.) dans le DICME est sous la direction de N.Vázquez Veiga. Un article lexicographique du DICME a les mêmes zones qu'un article du *Dictionnaire Explicatif et Combinatoire* [MEL'ČUK ET AL. 1984, 1988, 1992 et 1999], mais la zone sémantique est réduite à une étiquette sémantique du type 'action', 'événement', 'sentiment', 'artefact', etc. L'idée originelle d'un dictionnaire de collocations informatisé vient de A.Polguère, qui a déjà développé plusieurs milliers d'entrées pour le français dans son DiCo (voir [MEL'ČUK ET AL. 1995] et [POLGUÈRE, DANS CE VOLUME]). Dans notre projet, nous voulons entreprendre le premier dictionnaire de collocations pour l'espagnol.

<sup>2</sup>Pour le concept de *dérivation sémantique*, voir [POLGUÈRE DANS CE VOLUME]. *Grosso modo*, ce concept embrasse la synonymie, l'antonymie et les liens entre des noms qui désignent des participants dans la situation désignée par le mot-clé, comme par exemple, le nom d'agent, *criminel* et *crime*.

<sup>3</sup>Nous ne tenons pas compte ici des *expressions fusionnées* des FL syntagmatiques, comme Magn(*repas*)= //festin. La valeur de la FL ne se combine pas avec le mot-clé car *festin* signifie approximativement 'grand repas'.

<sup>4</sup>Il s'agit de la FL  $S_1(\textit{crime}) = \textit{criminel}$ .

<sup>5</sup>Notre représentation syntaxique des collocations ici n'est pas tout à fait conforme à la pratique de la TST. Dans ce cadre, dans le niveau syntaxique profond, les lexies correspondant aux valeurs des FL n'apparaissent pas, on y retrouve seulement le symbole de la FL. Ainsi, pour notre exemple, on aurait plutôt : ALIMENTATION -ATTR → Ver.

## Références

- [ALONSO RAMOS 1993] Alonso Ramos, Margarita (1993). *Las funciones léxicas en el modelo lexicográfico de Mel'čuk*, thèse de doctorat, U.N.E.D., Madrid.
- [ALONSO RAMOS/TUTIN 1996] Alonso Ramos, Margarita et Tutin, Agnès (1996). "A Classification and Description of Lexical Functions for the Analysis of their Combinations", in : [WANNER 1996], pp. 147-167.
- [IORDANSKAJA/MEL'ČUK 1995] Iordanskaja, Lidija et Mel'čuk, Igor (1995). "Traitement lexicographique de deux connecteurs textuels du français contemporain : *en fait* vs. *en réalité*", in : Hava Bat-Zeev Shyldkrot et Lucien Kupferman (éd.), *Tendances récentes en linguistique française et générale*, John Benjamins, Amsterdam/Philadelphia, pp. 211-236.
- [MEL'ČUK 1982] Mel'čuk, Igor (1982). "Lexical Functions in Lexicographic Description", in : *Proceedings of the Eighth Annual Meeting of the Berkeley Linguistics Society*, UCB, Berkeley, pp. 427-444.
- [MEL'ČUK 1996] Mel'čuk, Igor (1996). "Lexical Functions : A Tool for the Description of Lexical Relations in the Lexicon", in : [WANNER 1996], pp. 37-102.
- [MEL'ČUK ET AL. 1984] Mel'čuk, Igor, Arbatchewsky-Jumarie, N., Elnitsky, L., Iordanskaja, L. et Lesard, A. (1984). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques I*, Les Presses de l'Université de Montréal, Montréal.

- [MEL'ČUK ET AL. 1988] Mel'čuk, Igor, Arbatchewsky-Jumarie, N., Dagenais, L., Elnitsky, L., Iordanskaja, L., Lefebvre, M. N. et Mantha, S. (1988). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques II*, Les Presses de l'Université de Montréal : Montréal.
- [MEL'ČUK ET AL. 1992] Mel'čuk, Igor, Arbatchewsky-Jumarie, N., Iordanskaja, L. et Mantha, S. (1992) *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques III*, Les Presses de l'Université de Montréal, Montréal.
- [MEL'ČUK ET AL. 1999] Mel'čuk, Igor, Arbatchewsky-Jumarie, N., Iordanskaja, L., Mantha, S. et Polguère, A. (1999). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques IV*, Les Presses de l'Université de Montréal, Montréal.
- [MEL'ČUK ET AL. 1995] Mel'čuk, Igor, Clas, André et Polguère, Alain (1995). *Introduction à la lexicologie explicative et combinatoire*, Duculot, Louvain-la-Neuve.
- [ROBIN 1990] Robin, Jacques (1990), *Lexical Choice in Natural Language Generation*, Technical Report CUCS-040-090.
- [WANNER 1996] Wanner, Leo (éd.) (1996). *Lexical Functions in Lexicography and Natural Language Processing*, John Benjamins, Amsterdam/Philadelphia.
- [ZHOLKOVSKY/MEL'ČUK 1970] Zholkovsky, Alexander et Mel'čuk, Igor (1970 [1967]), "Sur la synthèse sémantique", in : *T. A. Informations*, 2, pp. 1-85.

