

Université de Montréal

**Représentations du sens lexical en lexicologie explicative et  
combinatoire :  
étude de trois formalisations spécifiques**

par

Mélissa Samson-Fauteux

Département de linguistique et de traduction

Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures  
en vue de l'obtention du grade de Maître ès sciences (M.Sc.)  
en linguistique

Août, 2010

© Mélissa Samson-Fauteux, 2010

Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé :

Représentations du sens lexical en lexicologie explicative et combinatoire :  
étude de trois formalisations spécifiques

Présenté par :

Mélissa Samson-Fauteux

évalué par un jury composé des personnes suivantes :

Louise Dagenais, président-rapporteur

Igor Mel'čuk, directeur de recherche

Jasmina Milićević, codirectrice

Pascale Lefrançois, membre du jury

**Mémoire accepté le : .....**

## Résumé

Cette étude porte sur les différentes formalisations du sens linguistique dans le cadre de la théorie Sens-Texte (TST) ; elle s'intéresse notamment à la représentation sémantique, la représentation la plus importante dans la modélisation du sens d'énoncés et d'unités lexicales de la langue. Cette étude a trois objectifs : premièrement, décrire trois formalisations de la représentation sémantique existant dans la TST – les réseaux sémantiques, les définitions lexicographiques et les définitions de la BDéf (des définitions très formalisées) ; deuxièmement, identifier les disparités et les décalages entre les formalisations choisies ; troisièmement, proposer des modifications aux formalisations existantes pour réduire au maximum ces disparités.

**Mots-clés :** Définition lexicographique, sémantique lexicale, formalisation de la sémantique, Lexicologie explicative et combinatoire, réseaux sémantiques.

## **Abstract**

This study explores the different formalisations of linguistic meaning as encountered in the Meaning-Text Theory (MST) framework. Its main interest is semantic representation, which is the most important representation in the modeling of the meaning of utterances and lexical units. This study has three objectives: first, describing three existing formalisations of semantic representation in the MST – semantic networks, lexicographical definitions, and BDéf definitions (definitions that are extremely formalized); second, identifying the discrepancies between these formalisations; finally, proposing modifications to these formalisations in order to reduce the disparities as much as possible.

**Keywords:** Lexicographical definition, lexical semantic, semantic formalization, Explanatory combinatorial lexicology, semantic networks.

# Table des matières

Résumé .....	iii
Abstract .....	iv
Table des matières .....	v
Liste des tableaux .....	ix
Liste des figures .....	ix
Liste des définitions .....	x
Liste des réseaux sémantiques .....	x
Liste des définitions BDéf complètes .....	xi
Liste des schémas de définitions .....	xi
Conventions d'écriture .....	xiii
Liste des abréviations récurrentes .....	xiv
<b>CHAPITRE 1 Problématique et cadre théorique .....</b>	<b>1</b>
1 Introduction .....	1
2 Notre cadre théorique .....	2
2.1 Le rôle du lexique .....	2
2.2 Les représentations .....	3
3 La représentation sémantique (RSém) .....	4
3.1 Les structures constituant la RSém .....	5
3.2 Description et délimitation de notre sujet d'étude : la RSém .....	6
4 Théorie du Métalangage Sémantique Naturel .....	8
4.1 Présentation générale .....	8
4.2 Quelques différences entre le MSN et la LEC .....	12
<b>CHAPITRE 2 Les formalisations de la RSém .....</b>	<b>16</b>
Introduction .....	16
1 Les différentes formalisations de la RSém .....	16
2 Description des trois formalisations choisies .....	19
2.1 Le réseau sémantique .....	23
2.2 La définition DEC .....	28
2.3 La BDéf .....	32
3 La structure communicative (SSém-Comm) .....	36

3.1	La thémativité.....	36
3.2	La présupposition.....	39
3.3	La locutionalité.....	39
4	Les étiquettes sémantiques.....	40
5	Questions théoriques non résolues.....	43
5.1	La représentation de la thémativité dans une RSém de lexie.....	43
5.2	Le rôle de la chronologie des faits dans l'ordre de lecture des composantes.....	43
5.3	Les conjonctions.....	45
5.4	Caractérisation de variables vs typage de variables.....	46
5.5	Les lexies à composante générique double.....	47
CHAPITRE 3 Les disparités dans les formalisations traitées.....		50
1	Introduction.....	50
2	Disparités entre les trois formalisations ayant trait à la structure sémantique (SSém)	
	51	
2.1	Composante générique et nœud communicativement dominant.....	51
2.2	Composantes périphériques.....	52
2.3	Explicitation des liens de polysémie.....	57
2.4	Contraintes sur les actants.....	58
2.5	Les connecteurs.....	60
2.6	Composantes faibles et optionnelles.....	60
3	Les disparités qui concernent la structure communicative (SSém-Comm).....	61
3.1	La thémativité.....	61
3.2	La locutionalité.....	62
3.3	La présupposition.....	62
4	Tableau synthèse des disparités entre les formalisations.....	62
CHAPITRE 4 Proposition d'améliorations à apporter aux trois formalisations de la RSém		
	64	
1	Introduction.....	64
2	Les schémas de définitions.....	65
2.1	Schéma proposé par Iordanskaja et Mel'čuk.....	66
2.2	Schéma proposé par Padučeva.....	67
2.3	Schéma proposé par Mel'čuk et Wanner.....	68

2.4	Schéma proposé par Milićević .....	69
2.5	Wierzbicka et le MSN .....	71
3	Modifications à apporter au réseau sémantique .....	72
3.1	Formalisations alternatives de réseau sémantique proposées par d'autres auteurs 73	
3.2	Proposition de modifications à apporter au réseau sémantique .....	74
3.3	Le rôle du schéma de définition .....	76
3.4	Les actants .....	77
3.5	Composante générique .....	78
3.6	Les liens de polysémie .....	79
3.7	La présupposition .....	80
3.8	La locutionnalité.....	80
3.9	Les composantes faibles et optionnelles .....	81
3.10	Organisation des blocs et liens entre ceux-ci .....	82
3.11	D'autres exemples de réseaux de lexies .....	84
4	Modifications à apporter à la définition DEC .....	85
4.1	Démarcation de la composante générique et du NDg .....	85
4.2	Contraintes sur les actants .....	87
4.3	Les connecteurs .....	88
4.4	Organisation en blocs définitionnels .....	89
4.5	Les liens de polysémie .....	90
4.6	La locutionnalité.....	90
4.7	Remarque sur la définition DEC .....	91
5	Modifications à apporter à la définition BDéf .....	91
5.1	Démarcation de la composante générique et du nœud communicativement dominant.....	91
5.2	Organisation en blocs définitionnels .....	91
5.3	Explicitation de l'ordre de lecture des blocs définitionnels .....	92
5.4	Composantes faibles et optionnelles .....	92
5.5	Les liens de polysémie .....	92
5.6	La présupposition .....	92
5.7	La locutionnalité.....	93

6 Améliorations à apporter aux formalismes concernant la SSém-Comm dans les schémas de définition.....	93
7 Tableau final.....	95
CHAPITRE 5 Conclusion.....	97
Bibliographie.....	xiv

## Liste des tableaux

Tableau I : Tableau résumant les éléments de formalisme associés à chacun des concepts lexicologiques analysés, et ce, pour chacune des trois formalisations traitées dans ce travail.....	63
Tableau II: Tableau résumant les éléments de formalisme proposés pour les concepts lexicologiques analysés dans les trois formalisations traitées.....	96

## Liste des figures

Figure 1 : Les représentations d'un énoncé dans la théorie Sens-Texte. ....	4
Figure 2: Liste des primitifs sémantiques proposée en 2007 par le MSN .....	10
Figure 3 : Définition de HANDLE de Wierzbicka .....	14
Figure 4 : Extrait de la définition de CUP .....	14
Figure 5: Représentation du schéma actanciel de TENIR-I.1 .....	24
Figure 6: Formalisme utilisé pour représenter la thémativité dans le réseau sémantique ....	25
Figure 7: Définition alternative #1 construite à partir du Réseau sémantique 2 .....	27
Figure 8: Définition alternative #2 construite à partir du Réseau sémantique 2 .....	28
Figure 9 : Présentation des différentes composantes d'une définition BDéf .....	35
Figure 10 : Extrait du DiCo montrant les colonnes "vocable", "numéro de vocable" et "étiquette" .....	41
Figure 11 : Extrait de la hiérarchie d'étiquettes sémantiques .....	42
Figure 12 : Le bloc /*Comportement*/ dans la définition BDéf de OURAGAN1 .....	43
Figure 13 : Le bloc de type d'actants dans la définition BDéf de PEIGNE1a.....	46
Figure 14 : Bloc élémentaire /*Caractérisation*/ dans la définition BDéf de OURAGAN1 ...	47
Figure 15 : Extrait de la définition BDéf de CHICHE1.1 .....	61
Figure 16 : Définition de REPROCHER1 : proposition de formalisation de Dostie <i>et al.</i> (1999) .....	73
Figure 17 : Définition de BOUILLOIREa : proposition de formalisation de Polguère (2002a) .....	74
Figure 18 : Composante générique du réseau sémantique de la lexie PLOMBIER-2 .....	79
Figure 19 : Illustration d'un en-tête de lien de polysémie .....	80
Figure 20 : Démarcation du cadre illocutoire dans nos réseaux sémantiques.....	81

Figure 21 : Format proposé appliqué au réseau sémantique de la RSém de la lexie BOUILLOIREa.....	84
Figure 22 : Format proposé appliqué au réseau sémantique de la RSém de la lexie OURAGANĪ .....	85
Figure 23 : Bloc central de la définition BDéf de OURAGANĪ .....	91

## Liste des définitions

Définition 1 : Définition de MANGER-1a .....	7
Définition 2 : Définition de FIANCÉ .....	7
Définition 3 : Définition DEC de TOMATEa .....	8
Définition 4 : Définition DEC de BERCEAUĪ.....	8
Définition 5: Définition de CUP de Wierzbicka .....	12
Définition 6 : Définition de CRAYON <sup>1</sup> du <i>Grand Robert</i> .....	17
Définition 7 : Définition de CRAYON du <i>Dictionnaire fondamental de la langue française</i> .17	
Définition 8 : Définition de CRAYONĪ.1 du DECFC-IV .....	18
Définition 9 : Définition DEC de OURAGANĪ.....	21
Définition 10 : Définition DEC de PEIGNE1a.....	27
Définition 11 : Définition DEC de SE BERCEĪ.1 .....	30
Définition 12 : Définition DEC de BOUILLOIREa .....	30
Définition 13 : Définition DEC de ÉTUDIANT .....	31
Définition 14 : Définition DEC de BERCEAUĪ.1 .....	32
Définition 15 : Définition DEC de <sup>Γ</sup> SE CASSER LE COU <sup>Γ</sup> .....	39
Définition 16 : Définition de BOJAT´SJAI.1a.....	67
Définition 17 : Définition de CONDAMNER <sup>1</sup> Ī.1 .....	71
Définition 18: Définition DEC de BOLa.....	90

## Liste des réseaux sémantiques

Réseau sémantique 1 : Réseau sémantique de la lexie OURAGANĪ .....	21
Réseau sémantique 2 : Réseau sémantique de PEIGNE1a .....	26
Réseau sémantique 3 : Réseau sémantique de la lexie OURAGANĪ avec indications temporelles .....	44

Réseau sémantique 4 : Proposition d'un nouveau format pour le réseau sémantique d'une RSém de lexie, appliqué à la lexie PEIGNE1a .....	75
Réseau sémantique 5 : Format proposé appliqué au réseau sémantique de la RSém de la lexie BERCEAUII.1.....	82

## Liste des définitions BDéf complètes

Définition BDéf 1 : Définition BDéf de OURAGANI .....	22
---	----

## Liste des schémas de définitions

Schéma de définition 1 : Schéma de définition de Iordanskaja & Mel'čuk (1990 : 17) .....	66
Schéma de définition 2 : Schéma de définition de Padučeva (1997 : 62).....	68
Schéma de définition 3 : Schéma de définition de Mel'čuk & Wanner (2001 : 15).....	68
Schéma de définition 4 : Schéma de définition de Milićević (2008 : 4).....	70
Schéma de définition 5 : Schéma de définition de Wierzbicka (1985 : 52).....	72

## Remerciements

Je voudrais remercier tous les gens qui m'ont permis de mener ce mémoire à terme.

D'abord, j'ai eu le privilège d'être dirigée par Igor Mel'čuk, qui est à la source de mon amour inconditionnel de la linguistique. M. Mel'čuk est la patience même et, tout au long de la rédaction de ce mémoire, il a su me guider de ses sages conseils et de ses pictogrammes évocateurs. J'aimerais le remercier pour sa disponibilité et pour m'avoir appris à toujours viser la rigueur et la vérité.

Je suis également reconnaissante à Jasmina Milićević, qui m'a codirigée. Elle a été d'une aide essentielle. Ces nombreuses relectures, son attention aux détails et ses judicieux conseils m'ont permis de compléter ce mémoire.

Merci également à Alain Polguère, pour m'avoir initiée à la recherche en lexicologie, pour le temps qu'il m'a consacré et pour m'avoir aidée à des moments où je devais faire des choix difficiles.

Surtout, à Ariel, qui m'a supportée (et endurée!) tout au long de ma maîtrise et qui a su glisser les mots d'encouragements exactement quand il le fallait : merci!

Un grand merci à Gilbert, qui a très gentiment révisé ce mémoire.

Finalement, je remercie le Département de Linguistique et de Traduction de l'Université de Montréal, qui m'a octroyé une bourse d'inscription ainsi qu'une bourse de rédaction pour ce mémoire.

## Conventions d'écriture

Voici les conventions d'écriture utilisées dans le cadre de ce travail pour les notions générales. Il ne s'agit aucunement des conventions propres aux différentes formalisations décrites.

- Lorsqu'une nouvelle notion est introduite, le nom de cette notion sera écrit en utilisant la police de caractères Helvetica, soit : **nouvelle notion**.
- Lorsqu'un terme est mentionné en tant que terme, on utilise l'italique : *terme*.
- Le soulignement simple sert à mettre une emphase : emphase.
- Le soulignement double est utilisé à la première mention d'une définition pour signifier que nous avons apporté des modifications à la définition citée : modification.
- Une composante sémantique, c'est-à-dire un sémantème ou une configuration de sémantèmes, est écrite entre guillemets simples : 'animal', 'émotion', 'tempête très violente'.
- Une unité lexicale, ou lexie, est écrite en petites majuscules : LEXIE.
- Une locution est délimitée par des demi-crochets supérieurs : 「PHRASÈME」.
- Les lexies doivent être désambiguïsées à l'aide de numéros distinctifs tirés d'un dictionnaire. Lorsque le dictionnaire utilisé est l'un des volumes du *Dictionnaire explicatif et combinatoire du français contemporain* publiés, nous l'écrivons : LEXIE1a ; lorsqu'il s'agit du *Grand Robert* (Rey, 2005), le numéro distinctif est précédé d'un tiret : LEXIE-1.
- Des lexies homonymes appartenant à des vocables différents sont distinguées par un chiffre en exposant et en caractères gras : LEXIE<sup>1</sup> vs LEXIE<sup>2</sup>.
- Une étiquette sémantique est écrite en police non proportionnelle : étiquette sémantique.
- Finalement, le nom d'un champ sémantique est écrit en majuscules : ANIMAUX, SENTIMENTS.

## Liste des abréviations récurrentes

### Notions :

CG : composante générique

DEC : Dictionnaire explicatif et combinatoire

LEC : Lexicologie explicative et combinatoire

ND : nœud communicativement dominant

NDg : nœud communicativement dominant global

RSém : Représentation sémantique

SSém : Structure sémantique

SSém-Comm : Structure communicative de la RSém

TST : Théorie Sens-Texte

### Ouvrages :

CONL : *Communicative organization in natural language*

DECFC (I-II-III-IV) : *Dictionnaire explicatif et combinatoire du français contemporain*  
(I-II-III-IV)

ILEC : *Introduction à la lexicologie explicative et combinatoire*

LAF : *Lexique actif du français*

# CHAPITRE 1      Problématique et cadre théorique

## 1      Introduction

La théorie Sens-Texte, TST pour la suite de ce travail, postule qu'une lexie ou une phrase peuvent être décrites à l'aide d'une série de représentations : une représentation sémantique, deux représentations syntaxiques, deux représentations morphologiques et deux représentations phonologiques. Chacune de ces représentations doit être formalisée pour permettre un travail linguistique concret. Notre mémoire porte sur la représentation sémantique (ou RSém) en particulier, la représentation qui cherche à modéliser le sens linguistique.

Il sera question de différents formats dans lesquels la représentation sémantique peut être décrite formellement dans le cadre de la TST. Ces formats seront appelés formalisations. Nous ferons d'abord l'examen critique de ces formalisations pour ensuite avancer des propositions en vue de les améliorer.

Notre exposé sera organisé de la façon suivante :

Dans un premier temps, dans ce chapitre, les notions nécessaires à ce travail, c'est-à-dire les notions de base de la TST ainsi que les notions qui se rapportent à la représentation sémantique en particulier, seront introduites. Aussi, nous introduirons la théorie du Métalangage Sémantique Naturel, une théorie qui présente des aspects différents de la TST, dont nous nous inspirerons dans nos propositions de formalisations.

Ensuite, dans le chapitre 2, trois formalisations de la représentation sémantique utilisée dans la TST seront étudiées en détail : les réseaux sémantiques, les définitions du DEC et les définitions de la BDéf, une base de données contenant des définitions extrêmement formalisées à des fins de recherches en linguistique computationnelle et en lexicologie, proposée dans le cadre d'un projet de description lexicographique mené à l'OLST (Observatoire de linguistique Sens-Texte) de l'Université de Montréal. Pour chacune des trois formalisations choisies, nous indiquerons le cadre dans lequel elles

sont employées (emploi pédagogique, recherche linguistique, traitement automatique du langage) et spécifierons leur formalisme.

Dans le chapitre 3, nous identifierons les différences entre les formalisations en question, c'est-à-dire les éléments qui soit sont présents dans une formalisation, mais pas dans une autre, soit ne sont pas équivalents dans au moins deux formalisations, et ce, sans que ces différences ne soient justifiées par des fondements théoriques ou des limites inhérentes aux formalisations elles-mêmes.

Dans le chapitre 4, des modifications aux formalisations existantes seront proposées afin de réduire au maximum leurs disparités.

Nous concluons dans le chapitre 5.

## 2 Notre cadre théorique

Dans ce chapitre, nous présenterons, dans la section 2, les notions de base nécessaires à notre exposé. Ensuite, dans la section 3, nous présenterons les éléments constituant la représentation sémantique (RSém), le sujet de notre étude. Puis, à la section 4, nous tracerons le portrait d'une autre théorie, celle du Métalangage Sémantique Naturel, comportant des éléments de représentation du sens que nous utiliserons dans nos propositions de formalisations dans le chapitre 4.

La lexicologie explicative et combinatoire, dorénavant LEC, est la composante lexicologique de la théorie Sens-Texte. Nous tenons pour acquis que le lecteur a une certaine connaissance de cette théorie. Pour une présentation plus complète de ses notions de base et de ses formalismes, nous renvoyons le lecteur à Mel'čuk (1997), à Polguère (1998) et à *l'Introduction à la lexicologie explicative et combinatoire*, désormais ILEC (Mel'čuk *et al.*, 1995) pour un tour d'horizon de la LEC.

### 2.1 Le rôle du lexique

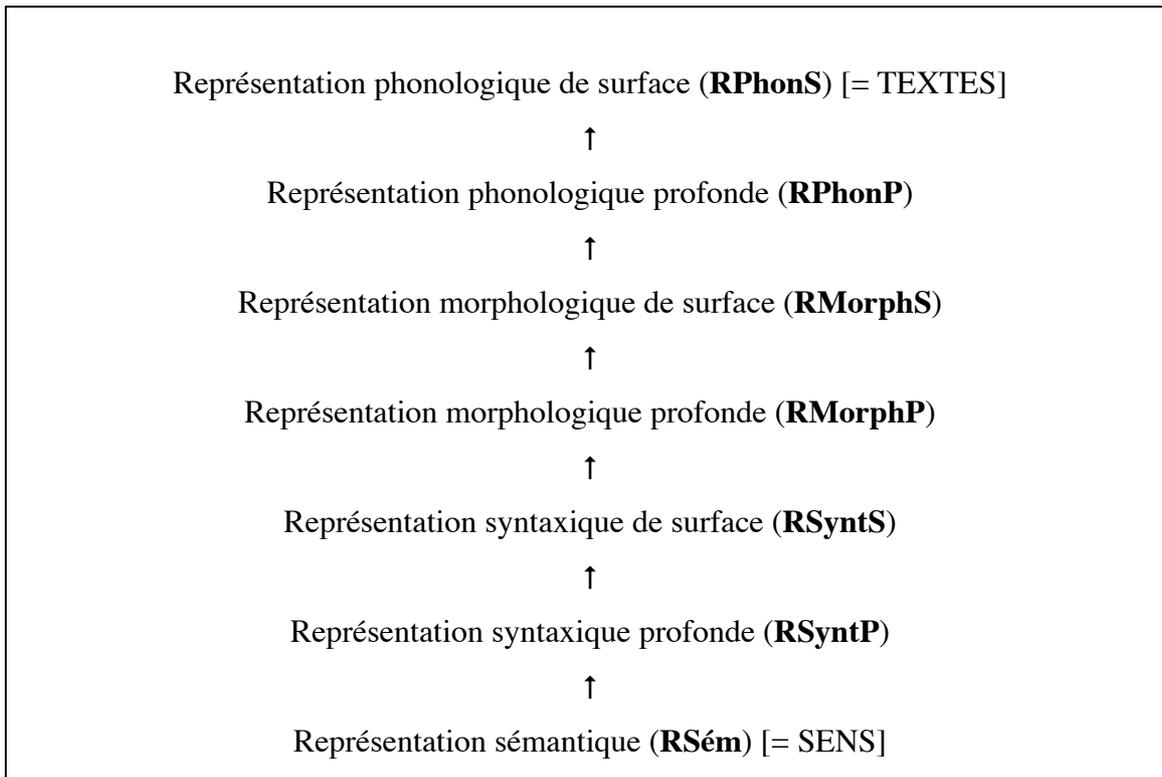
Selon la théorie Sens-Texte, le lexique occupe une place centrale dans la langue, le dictionnaire d'une langue constitue donc le cœur du modèle de cette langue (Polguère, 1998 : 17). Le lexique est composé de lexies : « On appellera *UNITÉ LEXICALE*, ou *LEXIE*, un élément de base du lexique. Chaque lexie de la langue est associée à un sens particulier, unique, définissable et à un ensemble de formes correspondant aux variantes

flexionnelles du signifiant de la lexie.» (Polguère, 2002b : 39). Ainsi, POMME<sup>1</sup>, CALMEMENT, POUR-**I.1**, 「TOMBER DANS LES POMMES<sup>7</sup> et DESSINER sont des lexies. On pense que le comportement d'une lexie dans un texte dépend en grande partie de son contenu sémantique (Mel'čuk *et al.*, 1995 : 73). Puisque le lexique revêt une telle importance dans l'étude de la sémantique d'une langue, il est tout naturel que la rédaction d'un dictionnaire constitue le centre des préoccupations dans le modèle Sens-Texte.

Ainsi, chaque langue doit posséder son Dictionnaire explicatif et combinatoire (DEC), un dictionnaire théorique dans lequel sont décrits, de façon exhaustive, le sens de chaque lexie de même que ses propriétés de combinatoire. La création d'un ouvrage comme celui-ci a débuté pour le français avec les quatre volumes du *Dictionnaire explicatif et combinatoire du français contemporain* (dorénavant DECFC) (Mel'čuk, 1984, 1988a, 1992; Mel'čuk *et al.*, 1999) parus à ce jour, ainsi qu'avec le *Lexique actif du français* (LAF) (Mel'čuk & Polguère, 2007), un dictionnaire de dérivations sémantiques et de collocations destiné aux apprenants du français.

## **2.2 Les représentations**

La TST s'intéresse avant tout à l'encodage du sens par le locuteur (synthèse) plutôt qu'à l'analyse du sens par le destinataire. La TST postule que la synthèse est en fait le seul processus réellement linguistique et qu'un phénomène linguistique se décrit à l'aide d'une succession de représentations qui permettent au locuteur de passer d'un Sens à un Texte. Voici les représentations qui permettent la production d'un énoncé linguistique complet (une phrase) :



**Figure 1 : Les représentations d'un énoncé dans la théorie Sens-Texte.**

Chacune de ces représentations comporte ses structures propres, qui sont spécifiées par des grammaires formelles. Pour chacune de ces structures, la TST propose un formalisme particulier (réseaux sémantiques, arbres de dépendance, chaînes morphologiques et phonologiques). L'élaboration du métalangage pour chaque représentation constitue une tâche essentielle dans la TST (Mel'čuk, 1988b : 51). L'autre volet indispensable d'un modèle Sens-Texte est l'encodage de règles de correspondance entre les niveaux de représentation adjacents qui permettent de passer du Sens (ce que le locuteur veut exprimer) au Texte (toutes les paraphrases permettant d'exprimer ce Sens).

### **3 La représentation sémantique (RSém)**

La LEC étudie tout ce qui se rapporte à la lexie : son sens, sa combinatoire lexicale, c'est-à-dire les collocations contrôlées par L, et sa combinatoire syntaxique, c'est-à-dire la sous-catégorisation ou le régime de L (les structures syntaxiques pour la réalisation des actants de L). Cependant, dans cet exposé, il sera question uniquement de l'aspect sémantique de la description d'une lexie, ce qui constitue un univers extrêmement riche et complexe en soi.

### 3.1 *Les structures constituant la RSém*

La RSém d'une phrase comporte les quatre structures suivantes (nous verrons que la RSém d'une lexie ne contient pas toutes ces structures) :

- 1) La **structure sémantique (SSém)** est formée de sémantèmes, c'est-à-dire de sens lexicaux de la langue considérée, et de relations prédicat ~ argument les reliant. La SSém est la **structure centrale** de la RSém : les autres structures se greffent à celle-ci. La SSém, représentée par le réseau sémantique, sera le sujet du chapitre 2, à la section 2.1.
- 2) La **structure sémantico-communicative (SSém-Comm)** se superpose à la SSém et permet au locuteur de restreindre le nombre de paraphrases qu'il peut produire à partir d'une certaine SSém – en fonction de ses choix communicatifs, voir Mel'čuk (2001). Elle représente la division de la SSém en aires communicatives marquées par des valeurs de huit oppositions communicatives :
  - 1) La thémativité : rhème / thème / spécifieur.
  - 2) Le donné/nouveau : donné / nouveau.
  - 3) La focalisation : focalisé / neutre.
  - 4) La perspective : arrière-plan / avant-plan / neutre.
  - 5) L'emphase : emphatique / neutre.
  - 6) La présupposition : posé / présupposé.
  - 7) L'unitarité : unitaire / articulé.
  - 8) La locutionnalité : communiqué / signalé / performé.

Les oppositions communicatives pertinentes pour notre exposé seront décrites dans le chapitre 2, à la section 3.
- 3) La **structure rhétorique (SRhét)** encode le registre et l'information stylistique au moyen de marques d'usage (poétique, humoristique, ironique, pathétique, etc.).
- 4) La **structure référentielle (SRéf)** lie les sémantèmes à leur référent, s'il y a lieu.

Les deux dernières structures sont très peu documentées dans la littérature de la TST. Leurs formalisations n'ont pas été définies à ce jour. Pour ces raisons, il n'en sera pas question dans le cadre de ce travail.

### 3.2 Description et délimitation de notre sujet d'étude : la RSém

La RSém modélise le sens de phrases et de lexies. Ces deux objets linguistiques ne se décrivent pas tout à fait de la même façon. La RSém d'une lexie ne possède pas de SRéf. Aussi, les RSém d'une lexie et d'une phrase ne contiennent pas nécessairement les mêmes données – les mêmes oppositions communicatives ne sont pas pertinentes pour la RSém d'une phrase et celle d'une lexie. Dans ce travail, nous traiterons uniquement de la formalisation de RSém de lexies, c'est-à-dire que nous traitons de la formalisation du sens d'expressions linguistiques comme MANGER-1a et POMME1 mais pas de la formalisation du sens d'expressions linguistiques comme *Jean mange une pomme*.

Il existe, en fait, deux types de lexies selon la structure de leur signifié : les lexies dénotant des objets sémantiques et les lexies dénotant des prédicats sémantiques, appelées **lexies prédictives**. Chaque lexie prédictive, comme MANGER-1a, dénote une situation qui présuppose des participants, dans ce cas-ci la personne ou l'animal qui mange et ce qui est mangé. Les sens qui dénotent des participants de la situation dénotée par une lexie L et qui s'expriment dans le texte sont des **actants** de cette lexie, représentés – dans le lexique – par des **positions actanciennes** de L « en attente » d'être remplies par des lexies. Les positions actanciennes sont instanciées par des variables ('X', 'Y', 'Z', etc.). Les objets sémantiques, quant à eux, sont des lexies qui ne présupposent pas de participants, et n'ont donc pas d'actants. Il s'agit de lexies du type SOLEIL, POMME, IGOR.

La description d'une lexie L, qu'il s'agisse de la dénotation d'un objet sémantique ou d'une lexie prédictive, se fait sous forme de définition verbale au sens large, c'est-à-dire sous la forme d'une paraphrase de L, une reformulation qui préserve le sens de L. La définition est composée de deux parties : le **défini** (ou *definiendum*) et le **définissant** (ou *definiens*). Ces deux expressions sont dans une relation d'égalité sémantique. Le défini d'une lexie prédictive est en fait la **forme propositionnelle** de la lexie décrite L, c'est-à-dire le sémantème correspondant à L, avec ses actants. Il s'agit de la partie à droite des deux-points :

**MANGER-1a** : 'X mange Y'

Le définissant est la RSém de la lexie. En conformité avec le modèle des définitions analytiques, utilisé depuis Aristote, le définissant doit être composé de sens plus simples

que celui de la lexie décrite. « Nous disons que la lexie  $L_1$  est *sémaniquement plus simple* que la lexie  $L$  si et seulement si on a besoin de  $L_1$  comme définissant de  $L$  alors qu'on ne peut utiliser  $L$  pour définir  $L_1$ . » (Mel'čuk *et al.*, 1995 : 80). Par exemple, on a besoin de 'meuble' pour définir 'berceau' (= 'meuble qui...'), mais on n'a pas besoin de 'berceau' pour définir 'meuble' ( $\neq$  'berceau qui...').

Le définissant est formé de deux blocs majeurs de composantes, soit la **composante générique** et les **composantes périphériques**, aussi appelées *genre prochain* et *différences spécifiques*. La composante générique est une paraphrase minimale du sens de la lexie définie, tandis que les composantes périphériques « caractérisent le sens de la lexie définie par rapport 1) à son genre prochain et 2) au sens de toutes les autres lexies de la langue ayant le même genre prochain » (Polguère, 2003a : 150). Voici quelques exemples de définitions paraphrastiques, dans lesquels le défini et le définissant sont réunis par le signe de l'égalité et dans lesquels les composantes génériques et périphériques sont identifiées :

**MANGER-1a**

'X mange Y' =

'L'être humain ou l'animal X [se nourrit]<sub>composante générique</sub> de Y, [qui est plus ou moins solide]<sub>composante périphérique-1</sub>, [en l'avalant]<sub>composante périphérique-2</sub> [(après l'avoir mâché)]<sub>composante périphérique-3</sub>'

**Définition 1 : Définition de MANGER-1a**

**FIANCÉ<sup>1</sup>**

'X est le ~ de Y' =

'[homme X]<sub>composante générique</sub> [tel qu'il s'est formellement engagé à épouser la femme Y]<sub>composante périphérique</sub>'

**Définition 2 : Définition de FIANCÉ**

<sup>1</sup> Le *Grand Robert de la langue française* compte une autre acception de la lexie FIANCÉ signifiant 'amant'. Si on tient compte de cette acception, la lexie définie ici serait plutôt FIANCÉa.

TOMATE<sub>a</sub> (DECFC-IV : 327)

‘Tomate’ =

‘[Unité de légumes<sub>I.a</sub>]<sub>composante générique</sub> [qui est de grosseur moyenne]<sub>composante périphérique-1</sub> et [de forme arrondie]<sub>composante périphérique-2</sub> et [dont la chair molle et la peau mince sont rouges<sub>I.1b</sub> (ou vertes<sup>1</sup><sub>I.1b</sub> quand elle n’est pas mûre)]<sub>composante périphérique-3</sub>’

**Définition 3 : Définition DEC de TOMATE<sub>a</sub>**

BERCEAU<sub>I</sub> (DECFC-IV : 114)

*Berceau de X* =

[Meuble]<sub>composante générique</sub> [destiné à ce qu’un bébé X y soit allongé pour y dormir]<sub>composante périphérique</sub> [(et tel qu’on peut y bercer<sub>I X</sub>)]<sub>composante périphérique</sub>

**Définition 4 : Définition DEC de BERCEAU<sub>I</sub>**

Rappelons que l’utilisation du soulignement double signifie que nous avons apporté des modifications aux définitions citées. Les parenthèses, que nous avons ajoutées à la définition de BERCEAU<sub>I</sub>, signifient qu’il s’agit d’une composante faible, une notion sur laquelle nous reviendrons à la section 2.2.2.1 du chapitre 2.

Il s’agit de définitions du type de celles exposées ci-dessus qui seront étudiées dans ce travail, mais ces définitions peuvent se présenter sous différentes formes. C’est ce que nous verrons dans les chapitres suivants.

## 4 Théorie du Métalangage Sémantique Naturel

### 4.1 Présentation générale

Nous présenterons d’abord les travaux d’une chercheuse offrant des formalisations différentes de celles traitées dans le cadre de ce mémoire, mais apparentées à ces dernières. Bien qu’il ne s’agisse pas d’une formalisation Sens-Texte, elle apportera des éléments intéressants à notre analyse et, surtout, certains de ses éléments de formalismes pourront être récupérés dans nos propositions d’améliorations aux formalisations.

La théorie du Métalangage Sémantique Naturel (*Natural Semantic Metalanguage*), désormais MSN, a été développée par Anna Wierzbicka (Wierzbicka, 1972, 1980, 1985, 1996; Goddard & Wierzbicka, 1994). Cette théorie ne sera pas décrite en détail ici ; les

bases de la théorie, ainsi que les références nécessaires sont disponibles sur le site du MSN (Goddard, 2008).

Comme dans la LEC, les définitions de Wierzbicka sont des « paraphrases réductives », c'est-à-dire qu'elles sont composées de sens plus simples que le sens de la lexie décrite. Aussi, elle évite l'usage de termes techniques et de symboles logiques. Cette recherche de la simplicité dans la description lexicographique a pour but de prévenir les cercles vicieux et l'opacité que l'on trouve dans certains dictionnaires.

La différence principale par rapport aux définitions de la LEC est que, selon le MSN, chacune des lexies utilisées dans une définition devrait correspondre à un sens indécomposable, soit à un **primitif sémantique**, pierre angulaire de la théorie de Wierzbicka. D'après Wierzbicka, les primitifs sémantiques devraient être exprimables dans toutes les langues, ils sont universels. Voici une liste complète des primitifs figurant sur le site Internet du MSN en 2007 :

Substantives:	I, YOU, SOMEONE, PEOPLE, SOMETHING/THING, BODY
Relational substantives:	KIND, PART
Determiners:	THIS, THE SAME, OTHER/ELSE
Quantifiers:	ONE, TWO, SOME, ALL, MUCH/MANY
Evaluators:	GOOD, BAD
Descriptors:	BIG, SMALL
Mental predicates:	THINK, KNOW, WANT, FEEL, SEE, HEAR
Speech:	SAY, WORDS, TRUE
Actions, events, movement, contact:	DO, HAPPEN, MOVE, TOUCH
Location, existence, possession, specification:	BE (SOMEWHERE), THERE IS, HAVE, BE (SOMEONE/SOMETHING)
Life and death:	LIVE, DIE
Time:	WHEN/TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT
Space:	WHERE/PLACE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE
“Logical” concepts:	NOT, MAYBE, CAN, BECAUSE, IF
Intensifier, augmentor:	VERY, MORE
Similarity:	LIKE

**Figure 2: Liste des primitifs sémantiques proposée en 2007 par le MSN**

Maintenant, présentons à titre d'exemple la définition de la lexie anglaise CUP ('tasse') par Wierzbicka :

## CUP

A KIND OF THING THAT PEOPLE MAKE IMAGINING THINGS OF THIS KIND PEOPLE WOULD SAY THESE THINGS ABOUT THEM:	
they are made for people to use repeatedly for drinking hot liquids from such as tea or coffee one person from one thing of this kind being able to put them down on something else	PURPOSE
they are made of something rigid, smooth and easy to wash which liquids can't go into or pass through and which doesn't break easily in contact with hot liquid	MATERIAL
they are rounded and open at the top so that one can drink easily from them by tipping the top part slightly towards the mouth without any of the liquid going outside where one doesn't want it to go	APPEARANCE -top
the bottom is the same shape as the sides and it is flat so that they are not more difficult to make than they have to be so that things of this kind can be put down on something else that is flat	-bottom
they cannot be much wider than they are high so that the liquid inside doesn't cease to be hot before one can drink it all they cannot be much higher than they are wide so that they don't overturn easily when one puts them down somewhere	-proportions
they have to be big enough to be able to have not less hot liquid in than a person would be expected to want to drink of that kind of liquid at one time they cannot be too big for people to be able to raise them easily to the mouth full of liquid, with one hand	SIZE
IMAGINING THINGS OF THIS KIND PEOPLE WOULD ALSO SAY THESE THINGS ABOUT THEM:	
people drink from them when sitting at a table	USE
they are made and used in groups of things which look the same so that people drinking from them can each have one of such things looking the same so that they look nice together	-sets
they are made and used together with some other things made of the same stuff made for putting these things on one thing for a person to drink from, on one thing for putting such things on so that people can raise the things to drink from to the mouth to drink a little and then put them down on those other things which can be put down on the table those other things are similar to them in some ways so that they look nice together	-saucers
those other things are made in such a way that there can be some parts of them all around the bottom of the things that people drink from so that if any liquid goes down over the top of the things people drink from it will come onto and remain on those other things and will not get elsewhere where one doesn't want it to go	

things of this kind are made of something thin so that they are nice to look at and to drink from and of something that one can't see through so that things made of it can have pictures and patterns on them making them nice to look at, such as china	MATERIAL
the sides are rounded, not flat so that all the sides look the same and one can't say where one side ends and another begins so that one can easily hold things of this kind around with a hand when they don't have any hot liquid inside	APPEARANCE -sides
the sides are not straight but roundish so that looked at from the side the opposite sides, or their upper parts, look like rounded lines whose middle parts are further away from one another than their tops or bottoms	
the bottom is smaller than the top so that if some of the liquid goes down along the outer surface of the thing one is drinking from it will come to the middle of the thing it is on and will not go outside that thing to where one doesn't want it to go	-proportions
things of this kind have a thin rounded part which sticks out from one side which doesn't get hot because it is not in contact with the hot liquid inside this part is attached at its top and its bottom to one side so that by holding it one can prevent one part of the top from being above the others when one is not drinking and so to prevent any of the liquid from getting out over the top and going where one doesn't want it to go	-handle
this part has to be big enough and to stick out far enough for people to be able to raise things of this kind full of liquid for a short time holding that part with a thumb and two fingers without any parts of the hand touching the sides, which are hot it can't be much bigger or stick out much further than necessary for that	

**Définition 5: Définition de CUP de Wierzbicka**

Ainsi, les définitions de Wierzbicka se démarquent par l'utilisation d'un langage naturel et accessible, mais aussi, par leur longueur proportionnelle à la complexité sémantique des lexies décrites. Les indications dans la colonne de droite sont des en-têtes qui caractérisent les blocs de sens. La casse (majuscule ou minuscule) détermine le niveau d'importance du bloc. Aussi, Wierzbicka utilise les espaces entre les blocs de composantes et les différentes casses dans les en-têtes de blocs pour marquer respectivement la proximité des composantes de sens et leur importance dans la définition.

#### 4.2 *Quelques différences entre le MSN et la LEC*

Premièrement, la LEC utilise aussi le concept de primitifs sémantiques, mais opte pour la définition par des sémantèmes immédiatement plus simples que le sens de la lexie décrite plutôt que sur la recherche des primitifs sémantiques. On juge qu'il n'est pas

possible d'établir la liste des primitifs avant d'avoir décrit tout le lexique. Le choix des sémantèmes utilisés dans une définition de la LEC est dicté par la règle du bloc maximal :

La définition de la lexie vedette L doit être faite par des lexies  $L_1, L_2, \dots, L_n$  telles qu'aucune configuration de ces lexies ne peut être remplacée par une seule lexie sémantiquement équivalente (Mel'čuk *et al.*, 1995 : 84).

Cette règle contraint les lexicographes à toujours utiliser, dans une définition lexicographique, le sémantème le plus englobant possible. Elle est justifiée par le caractère arbitraire du choix d'un sémantème qui ne serait ni le plus englobant possible ni un primitif sémantique. Ainsi, la LEC postule que si l'on décrit toutes les lexies d'une langue avec des unités toujours immédiatement plus simples que la lexie décrite, puis que l'on décompose à leur tour ces lexies plus simples et que l'on continue ainsi jusqu'à ce que toutes les lexies d'une langue soient décrites, l'ensemble des primitifs sémantiques finira certainement par être découvert.

Du côté de Wierzbicka, on permet parfois l'usage de lexies intermédiaires préalablement décrites, en particulier pour la description d'artefacts et d'« espèces naturelles », mais on spécifie que ces lexies ne doivent pas faire partie du même domaine de spécialité que la lexie décrite (Wierzbicka, 1985 : 45). Ainsi, si l'on définit des lexies appartenant au champ sémantique des ANIMAUX, on doit éviter les sémantèmes 'patte', 'museau', 'niche', etc. qui sont directement liés au domaine des animaux.

Wierzbicka exprime tout de même des réserves quant à cette façon de faire et tente de recourir le moins possible à des sémantèmes intermédiaires dans ses définitions. Elle considère que, dans les faits, il s'agit souvent d'un « alourdissement » sémantique. Ainsi, dans la définition de CUP, en décrivant la poignée ('handle') sans utiliser 'handle' comme elle le fait ci-dessus, Wierzbicka dit tout ce qu'elle doit dire sur la poignée d'une tasse et rien de plus (voir le bloc « handle » de la définition de CUP). Wierzbicka affirme que si elle utilisait le sémantème 'handle', que l'on peut décomposer comme dans la définition de la Figure 3 ci-dessous, la définition serait alors redondante puisque des composantes comme 'object made by people for people to use by moving them in some ways' sont déjà exprimées ailleurs dans la définition de CUP.

**HANDLE**

a part of objects of any kind made by people for people to use by moving them in some ways which sticks out far enough for a person to be able to hold it easily with a hand or hands so that by holding it one can cause the whole thing or some other part of it to move in the way one wants them to move

Figure 3 : Définition de HANDLE de Wierzbicka

Dans les faits, Wierzbicka utilise beaucoup de sémantèmes qui ne sont pas des primitifs. Dans cet extrait de la définition de CUP, seulement le tiers des sémantèmes employés sont des primitifs sémantiques, ce sont les sémantèmes en caractères gras :

**things of this kind have a thin rounded part which sticks out from one side which doesn't get hot because *it is not in contact with the hot liquid inside* **this part is attached at its top and its bottom to one side** so that by holding it one can prevent one part of the top from being above the others when one is not drinking and so to prevent any of the liquid from getting out over the top and going where one doesn't want it to go **this part has to be big enough and to stick out far enough for people to be able to raise things of this kind full of liquid for a short time** holding that part with a thumb and two fingers without any part of the hand touching the sides, which are hot **it can't be much bigger or stick out much further than necessary for that****

Figure 4 : Extrait de la définition de CUP

Les lexies en italique gras sont des pronoms qui ne font pas partie de la liste des primitifs établie par le MSN. Wierzbicka explique qu'elle a recours à des pronoms pour que les définitions soient moins lourdes et plus naturelles. Les lexies en soulignement pointillé sont des lexies qui combinent deux primitifs sémantiques qui, lorsqu'utilisés séparément, font une phrase agrammaticale en anglais, par exemple *more big* doit être dit *bigger* en anglais. Les lexies soulignées sont des lexies dont le sens est complexe, ce ne sont pas des primitifs. Finalement, les lexies en gris sont sémantiquement vides (sauf pour les pronoms sémantiquement vides qui sont déjà identifiés comme pronoms).

**primitifs*****pronoms***

*composés de primitifs, pas grammatical si séparés*

*pas des primitifs*

*lexies vides : par ex., prépositions régies*

On pourrait penser que les consignes de rédaction du MSN devraient être précisées. Peut-être s'agirait-il d'exiger de décomposer seulement les éléments d'une définition qui appartiennent au même champ sémantique que la lexie décrite puisque le choix de

lexies intermédiaires comporte toujours sa part d'arbitraire. Pour cette raison, nous respecterons la règle du bloc maximal de la LEC tout en nous inspirant de la structure des définitions de Wierzbicka.

Wierzbicka nous influence sur au moins trois points. D'abord, l'idée d'utiliser des en-têtes pour identifier chacun des rôles informationnels dans une définition provient des définitions de Wierzbicka. Deuxièmement, comme cette dernière, nous sommes d'avis que les blocs définitionnels ne se trouvent pas tous au même niveau et qu'il est pertinent d'avoir recours à des blocs et des sous-blocs, là où Wierzbicka a recours à des en-têtes et des « sous-en-têtes », en se servant des casses différentes. Troisièmement, nous aimerions reproduire, à notre façon, l'idée de représenter visuellement la distance entre chacune des composantes sémantiques.

## CHAPITRE 2 Les formalisations de la RSém

### Introduction

Ce chapitre est organisé comme suit : nous présenterons, dans la section 1, les différents types de formalisations de la RSém qui existent dans la TST. Ensuite, dans la section 2, nous décrirons en détail les formalismes des trois formalisations étudiées (le réseau sémantique, la définition DEC et la définition BDéf). Nous exposerons différentes notions concernant la structure communicative à la section 3 et, à la section 4, nous présenterons la notion d'étiquette sémantique, un élément important de notre formalisation qui permet de classer les composantes génériques des lexies. Nous concluons ce chapitre en portant à l'attention du lecteur quelques questions théoriques soulevées par ce travail (section 5).

### 1 Les différentes formalisations de la RSém

Il existe différentes motivations pour une formalisation de RSém. Évidemment, la motivation influence grandement les éléments que contient la formalisation et la forme que présentent ces éléments. Une lexie ne sera pas définie de la même manière pour un apprenant d'une langue que pour une application informatique qui manipule le langage naturel de façon automatique. Pour le moment, nous connaissons deux types de formalisations de RSém : les définitions lexicographiques et les réseaux sémantiques. Ces deux types sont subdivisés en sous-types en fonction de l'utilisateur et de la tâche qu'elles visent à accomplir.

- 1) Les définitions verbales ou langagières :
  - a. Les définitions grand public : Ces définitions utilisent la langue naturelle, sous forme contrôlée – en restreignant le choix des lexies employées dans les définitions – ou non. Ce sont les définitions que l'on retrouve dans les dictionnaires d'usage courants (par ex. les définitions du *Petit Robert* (Rey-Debove, & Rey, 2010) ou du *Grand Robert* (Rey, 2005)).

CRAYON<sup>1</sup> (Rey, 2005)

2. Instrument servant à écrire, formé d'une petite baguette, généralement en bois, servant de gaine à une longue mine.

**Définition 6 : Définition de CRAYON<sup>1</sup> du *Grand Robert***

- b. Les définitions d'apprentissage : Ces définitions sont destinées aux apprenants d'une langue. Elles utilisent la langue naturelle, contrôlée ou non. Elles sont caractérisées par une simplification de leur forme et de leur contenu (par ex. les définitions du *Dire Autrement* (Hamel & Milićević, 2007; Milićević & Hamel, 2007; Milićević, 2008, 2009) ou du *Dictionnaire fondamental de la langue française* (Gougenheim, 1958).

CRAYON (*Id.*: 79)

instrument qui sert à écrire et à dessiner

**Définition 7 : Définition de CRAYON du *Dictionnaire fondamental de la langue française***

On remarque que la définition est simple et très courte. Elle n'est pas complète, elle vise plutôt la clarté et l'accessibilité. Cette définition est destinée à des apprenants débutants.

L'étude des formalismes à adopter pour les définitions destinées aux apprenants et des données lexicographiques à y inclure constitue un domaine très intéressant en soi, que nous ne pourrions pas aborder ici. Nous renvoyons le lecteur, entre autres, à Rundell (1998) et Lew & Dziemianko (2006).

- c. Les définitions destinées aux linguistes : Ces définitions s'adressent aux chercheurs en linguistique. Elles prennent également la forme d'une définition « langagière », mais sous forme strictement contrôlée. Elles visent la précision plutôt que la clarté (par ex. les définitions des DECFC I-II-III-IV).

CRAYON **I.1** (DECFC-IV : 201)

Crayon Z de X = Instrument<sub>1,1</sub> destiné à ce qu'une personne X l'utilise pour écrire ou pour dessiner **I.1,2a,2c** - bâtonnet en bois contenant au milieu une tige de substance de couleur Z qui laisse des traces sur la surface sur laquelle X écrit ou dessine **I.1,2a,2c**.

**Définition 8 : Définition de CRAYON I.1 du DECFC-IV**

On remarque que certaines des lexies au sein de la définition sont désambiguïsées, que les actants sont instanciés par des variables et que la définition est relativement longue. Ceci est une définition DEC de complexité moyenne.

2) Les réseaux sémantiques :

- a. Les réseaux sémantiques destinés aux linguistes : Ils sont très formalisés, mais demeurent lisibles pour des linguistes entraînés. C'est le type de réseaux dont il sera question dans le cadre de ce travail. Ils seront décrits à la section 2.1.
- b. Les réseaux sémantiques destinés au traitement automatique du langage : Ils doivent être formalisés au maximum pour être lus et manipulés de façon automatique.

Selon la LEC, la formalisation « ultime » du sens, tant de la lexie que de la phrase, devrait se faire sous forme de réseau sémantique. Cependant, pour la description du sens de lexies, la définition lexicographique est aussi utilisée parce qu'elle est plus facile à lire et permet aux linguistes d'utiliser leur intuition linguistique.

Plusieurs formalisations de définitions ont été développées dans le cadre de la TST afin de satisfaire des besoins de représentation qui ne sont pas toujours les mêmes. Il existe, entre autres, le format BDéf (Altman & Polguère, 2003) et la formalisation des définitions du *Dire Autrement* (Milićević, 2008), destinées aux apprenants. Ces deux derniers projets ont été réalisés d'abord à partir des définitions des DECFC. On peut aussi mentionner Padučeva avec la base de données des verbes russes « Lexicographer » (Padučeva & Kustova, 2004), voir Padučeva (2009) pour un extrait traduit en anglais.

## 2 Description des trois formalisations choisies

Puisque nous traitons de formalismes utilisés dans les formalisations considérées, il nous est nécessaire de définir une terminologie pour parler desdits formalismes. Ainsi, nous dirons que chaque *concept lexicologique* est représenté par un *élément de formalisme*. Par exemple, dans la LEC, le concept lexicologique de position actancielle est représenté par des variables ; l'utilisation de ces variables est donc un élément de formalisme. Aussi, le concept lexicologique de composante optionnelle est représenté par l'utilisation de parenthèses, qui constituent également un élément de formalisme.

Comme nous l'avons déjà dit, nous comparerons les formalismes utilisés dans les réseaux sémantiques, les définitions du DEC et les définitions de la BDéf. Nous avons choisi le réseau sémantique et la définition DEC parce qu'il s'agit des formalisations de base en TST. Nous avons choisi de traiter de la BDéf parce qu'elle propose de faire un pont entre les deux autres formalisations, en offrant un compromis entre, d'une part, la représentation extrêmement formelle du réseau sémantique et, de l'autre, la linéarité et la lisibilité de la définition lexicographique. Voici une liste des éléments de formalisme communs aux trois formalisations étudiées. Puisque chacun de ces éléments de formalisme est représenté de façon satisfaisante et équivalente d'une formalisation à l'autre, nous ne reviendrons pas sur ces éléments dans le cadre de ce travail.

- Le défini est présenté sous la forme propositionnelle et les positions actanciennes de la lexie L sont représentées par des variables. Par exemple, le défini de la lexie MANGER-**1a** est 'X mange-**1a** Y', dans lequel 'X' et 'Y' représentent ses deux positions actanciennes.
- Les lexies utilisées dans la représentation du sens de L sont désambiguïsées au moyen de numéros distinctifs. Par exemple, dans la définition de TOMATE<sub>a</sub>, le sémantème correspondant à la lexie LÉGUMES**I.a** est désambiguïsé des autres lexies LÉGUME contenues dans le même vocable :

'Tomate' = 'Unité de légumes**I.a** [...]'

- Les positions actanciennes sont typées, mais le typage n'est pas nécessairement explicite. Par exemple, dans la définition de MANGER-**1a** (voir la Définition 1 à la page 7), on spécifiera que 'X' est un être humain ou un animal, mais les

sémantèmes ‘être humain’ et ‘animal’ ne se distinguent pas des autres sémantèmes de la définition par un élément de formalisme particulier.

- La présupposition est identifiée, lorsqu’elle est présente. Nous reviendrons sur les formalismes impliqués par la présupposition.

Les concepts lexicographiques de la RSém qui ne sont pas mentionnés ici ne le sont pas parce que :

- 1) ils ne sont pas représentés du tout dans au moins une des formalisations décrites ;
- 2) le formalisme à utiliser pour les décrire n’a pas été défini.

Ces concepts lexicographiques seront traités aux chapitres 3 et 4.

Pour amorcer la présentation, nous présenterons un exemple de chacune des formalisations traitées. Voici, d’abord, le réseau sémantique de la lexie OURAGANI [*L’ouragan se déchaîne sur la mer des Caraïbes.*] :



Et finalement, voici la définition BDéf de cette même lexie OURAGANI :

<b>1: tempête#I.violent.très</b>	
<pre> /*Comportement*/ 2: *1 _se_former_ dans X 3: *1 _avoir_lieu_ dans Y 4: *1 _se_déplacer_ vers W  /*Caractérisation.X [présupposé]*/ 5: X être {région_tropicale} de océan#I1 <u>Atlantique</u>  /*Effet_habituel*/ 6: *1 tendre à *7 7: *1 avoir {effet}.destructeur sur Z  /*Localisation_spatiale.Z [présupposé]*/ 8: Z _se_trouver_ dans Y </pre>	
<b>Typage</b>	X: Y: espace <u>géographique</u> Z: W: <u>espace géographique</u>
<b>Rel_Sém</b>	

**Définition BDéf 1 : Définition BDéf de OURAGANI**

Nous pouvons maintenant décrire le formalisme de chacune des formalisations : le réseau sémantique, la définition DEC et la définition BDéf.

## 2.1 Le réseau sémantique

### 2.1.1 Caractérisation du réseau sémantique

#### 1) À qui s'adresse-t-il ?

Le réseau sémantique sert principalement à la recherche linguistique. Il aide le lexicographe à rédiger des définitions cohérentes et le lexicologue à réfléchir sur la structuration du sens des lexies.

Le réseau sémantique constitue aussi un outil privilégié pour le traitement automatique du langage, mais, afin d'être lu par un programme informatique, le réseau doit être traduit dans un format « à plat », linéaire. Cependant, notons que cette linéarité n'a aucune valeur organisationnelle et que l'ordre de chacune des composantes est aléatoire.

#### 2) À quel type d'entité linguistique s'applique-t-il ?

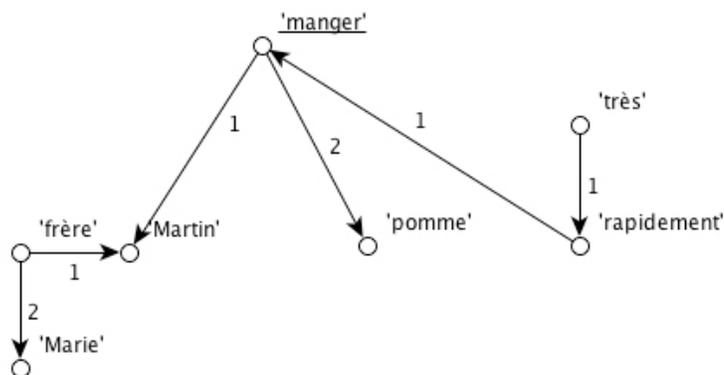
Un réseau sémantique peut représenter autant le sens d'une lexie que celui d'une phrase.

### 2.1.2 Description du réseau sémantique

Voici pour la phrase (1) un réseau sémantique en (2) :

(1) *Martin, le frère de Marie, mange une pomme très rapidement.*

(2)<sup>2</sup>



Le réseau sémantique est composé de nœuds portant des sémantèmes. Un sémantème associé à un référent donné n'est représenté qu'une seule fois dans le réseau. Des arcs

<sup>2</sup> Il s'agit d'un réseau simplifié pour illustrer ce qu'est un réseau sémantique. Un réseau complet contiendrait des indications de temps et de nombre et ses sémantèmes seraient désambiguïsés.

munis de flèches lient ces nœuds entre eux. Un arc représente une relation de dépendance sémantique (de type prédicat ~ argument) entre les sémantèmes. Lorsqu'il est question de la relation de dépendance sémantique, on parle de *prédicats* et de leurs *arguments*, mais, lorsqu'il s'agit d'une lexie<sup>3</sup> et des lexies qui en dépendent sémantiquement, on parle plutôt d'une lexie et de ses *actants sémantiques*. Le sémantème 'S', correspondant à une lexie prédicative L, est le *gouverneur sémantique* de ses actants sémantiques. Donc, le sémantème 'manger-1a' représenté ci-dessus, qui correspond à la lexie MANGER-1a, a deux actants sémantiques, 'Martin' et 'pomme'.

Les arcs portent des numéros actanciels, qui expriment la position de chacun des actants sémantiques dans le schéma actanciel de leur gouverneur sémantique. Par exemple, 'tenir-I.1' [*Jean tient le peigne dans sa main.*] est un prédicat à trois arguments : celui qui tient quelque chose, ce qu'il tient et ce avec quoi il le tient. Dans la définition de la lexie TENIR-I.1, ses arguments sont instanciés par des variables : 'X', 'Y' et 'Z'.

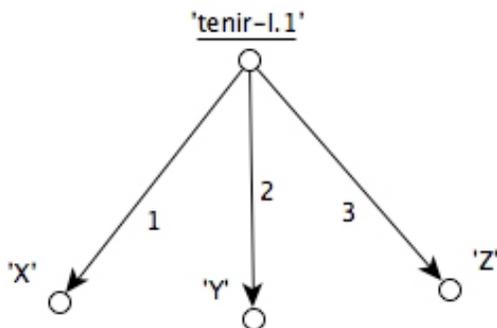


Figure 5: Représentation du schéma actanciel de TENIR-I.1

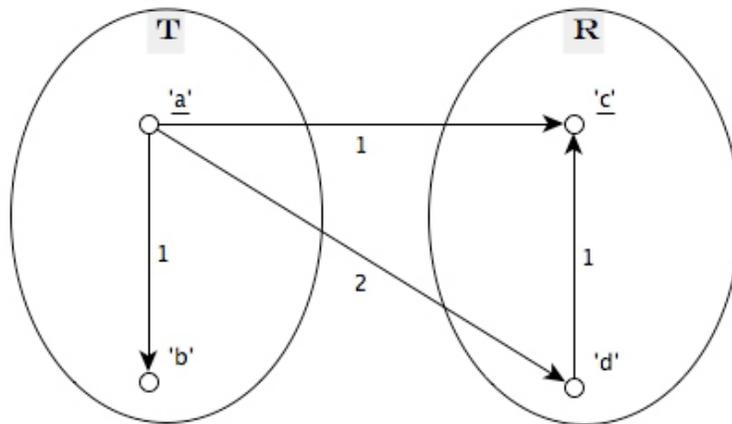
Notons que lorsqu'une définition contient un participant qui n'est pas un actant, parce qu'il ne s'exprime pas en surface, ce participant est représenté par une lettre grecque : 'α', 'β', etc.

Un réseau (ou un sous-réseau) sémantique a toujours un nœud communicativement dominant (ND), c'est-à-dire un nœud qui « résume » son contenu. Un sémantème S2 qui résume le sens de S1 est un sémantème contenu dans la décomposition sémantique de S1.

<sup>3</sup> Il s'agit en fait de sens de lexie, mais nous nous permettons ce raccourci, pour simplifier la présentation.

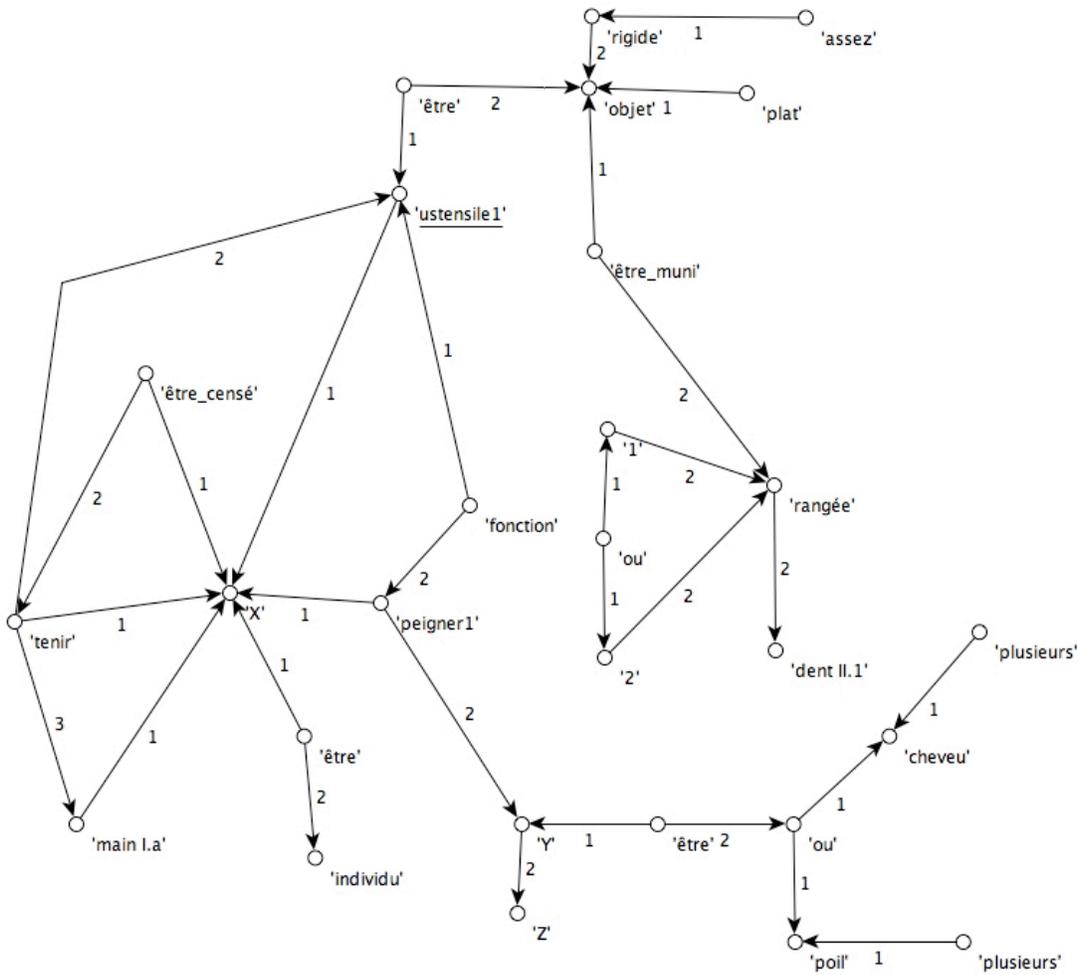
S2 est un sens hyperonymique de S1 : la lexie correspondant à S2 pourrait être utilisée à la place de la lexie correspondant à S1 dans un énoncé sans que l'on déforme le sens véhiculé – même si l'on en perd une partie plus ou moins importante. Dans le cas du ND, il s'agit toujours d'un seul sémantème et non d'une configuration de sémantèmes. Celui-ci est indiqué par un soulignement (Polguère, 2002a : 12). En fait, on parle de *nœud communicativement dominant* (ND) pour référer au nœud qui résume le sens d'un sous-réseau et de *nœud communicativement dominant global* (NDg), lorsque l'on parle du nœud qui résume le sens d'un réseau entier.

La SSém-Comm se superpose au réseau sémantique en y délimitant des aires communicatives. Dans le réseau présenté à la Figure 6 ci-dessous, les nœuds 'a', 'b', d'une part, et 'c', 'd', de l'autre, appartiennent à des aires communicatives différentes.



**Figure 6: Formalisme utilisé pour représenter la thémativité dans le réseau sémantique**

À partir de maintenant, nous ne parlerons de réseaux sémantiques que pour désigner une définition de lexie représentée sous forme de réseau sémantique. Voici, pour poursuivre notre exposé, le réseau sémantique qui présente la SSém de la lexie PEIGNE1a.



### Réseau sémantique 2 : Réseau sémantique de PEIGNE1a

L'utilisation du réseau sémantique est avantageuse puisqu'elle permet de montrer explicitement les propriétés formelles des éléments qui le composent (prédicat, objet sémantique). Il facilite la visualisation des liens de dépendance sémantique entre les sémantèmes. Par contre, l'interprétation d'un réseau sémantique ne va pas de soi, en particulier pour ce qui est de l'ordre et de la direction de la lecture. Il n'est pas évident que, à partir du Réseau sémantique 2, un lecteur reconstruirait la Définition 10 ci-dessous, qui a pourtant servi à le construire.

PEIGNE**1a** (DECFC-IV : 281):

*Peigne de X [pour peigner Y de Z] = Ustensile**1** de toilette destiné à ce que l'individu X peigne**1** les cheveux ou le poil Y de Z – objet plat assez rigide censé être tenu à la main**I.a** par X, et muni d'une ou de deux rangée(s) de dents**II.1**.<sup>4</sup>*

**Définition 10 : Définition DEC de PEIGNE1a**

Un réseau comme le Réseau sémantique 2 peut être lu de dizaines de façons différentes. Bien sûr, les mêmes sémantèmes se trouveront dans chacune de ces lectures, ce sont plutôt les rapports entre les composantes qui changent – deux composantes sont plus ou moins éloignées l'une de l'autre, ou encore une composante vient avant ou après une autre –, ce qui modifie la logique de présentation et la compréhension du sens de la lexie décrite. Ainsi, on entrevoit que le sens d'une lexie doit être plus structuré que ne l'est le Réseau sémantique 2 présenté ci-dessus.

Voici deux exemples de lectures possibles du Réseau sémantique 2 qui ne correspondent pas à la Définition 10, qui, rappelons-le, a servi à écrire ce réseau :

PEIGNE**1a**

ustensile**1**

censé être tenu à la main**I.a** par X

qui est un objet muni d'une ou plusieurs rangées de dents**II.1**

dont la fonction est que X, qui est un individu, peigne**1** les cheveux ou le poil Y

qui est un objet assez rigide

qui est un objet plat

**Figure 7: Définition alternative #1 construite à partir du Réseau sémantique 2**

<sup>4</sup> Les définitions d'artefacts comme celles-ci ont un formalisme particulier et ont deux composantes génériques, ici 'ustensile**1**' et 'objet'. Nous reviendrons sur ce type de définition à la section 5.5 du chapitre 2.

<p>PEIGNE<b>1a</b></p> <p>ustensile<b>1</b></p> <p>de X</p> <p>qui est un individu</p> <p>qui est censé tenir l'ustensile à la main<b>1a</b></p> <p>qui peigne<b>1</b> les cheveux ou le poil Y, ce qui est la fonction de l'ustensile</p> <p>qui est un objet muni d'une ou plusieurs rangées de dents<b>II.1</b></p> <p>qui est un objet assez rigide et plat</p>
---

**Figure 8: Définition alternative #2 construite à partir du Réseau sémantique 2**

En fait, à l'exception du soulignement du NDg, le seul formalisme existant dans le réseau sémantique pour distinguer des composantes de statut différent est le sous-réseau, c'est-à-dire un encerclement d'une configuration de sémantèmes ayant les mêmes propriétés. On utilise les sous-réseaux pour représenter les oppositions communicatives, mais aussi pour représenter les composantes faibles ou optionnelles d'une définition (ces composantes seront définies à la section 2.2.2.1).

## 2.2 La définition DEC

### 2.2.1 Caractérisation de la définition DEC

#### 1) À qui s'adresse-t-elle ?

La définition lexicographique DEC est destinée aux linguistes qui veulent réfléchir sur le sens lexical. Lorsqu'on lui apporte quelques simplifications, la définition DEC peut devenir une définition pour les apprenants (à l'instar de celles mentionnées au chapitre 3, à la section 1), comme le démontre le travail fait dans le cadre du dictionnaire *Dire Autrement*.

#### 2) À quel type d'entité linguistique s'applique-t-elle ?

En règle générale, on ne définit que des lexies. Toutefois, il est possible de composer une paraphrase « à la DEC » pour une phrase dans laquelle les lexies seraient désambiguïsées et dans laquelle certains phénomènes comme la présupposition seraient marqués. En fait, Padučeva (1997) construit des définitions de phrases du type *John killed a fly*, dans lesquelles le sens de phrases est décomposé en sens plus simples.

### 2.2.2 Description de la définition DEC

Le terme *définition DEC* est employé pour parler des définitions de la LEC que l'on retrouve dans les DECFC. Les auteurs du DECFC ont jugé la définition lexicographique plus adéquate que le réseau sémantique pour cet ouvrage parce qu'elle est beaucoup plus facile à interpréter et permet à l'intuition linguistique de s'exprimer. C'est surtout son caractère verbal, c'est-à-dire langagier, qui facilite son interprétation. Ainsi, la définition lexicographique de la LEC se rapproche des définitions des dictionnaires d'usage parce qu'elle est rédigée « en prose » dans une langue naturelle. Cependant, dans les DEC, cette langue est très contrôlée, ce qui signifie, entre autres, que chaque lexie employée dans des définitions doit être désambiguïsée et que la synonymie, au sein de ces définitions, est limitée au maximum. Ceci vise à prévenir les définitions ambiguës (qui correspondent à plusieurs interprétations possibles) ou circulaires (qui contiennent des sens qui sont eux-mêmes définis par la lexie décrite). Le fait d'éviter l'utilisation de lexies synonymes permet également de repérer des régularités d'une définition à une autre. D'autres dictionnaires, mais ils sont assez rares, se servent d'un vocabulaire contrôlé, citons Wierzbicka (1972, 1985, 1987, 1996), dont la théorie a été présentée à la section 4 du chapitre 1, le *Longman dictionary of contemporary English* ou *LDOCE* (Fox, 2007), qui utilise un vocabulaire contrôlé de 2000 mots, et le *Dictionnaire du français fondamental* (Gougenheim, 1958), qui lui a recours à un vocabulaire de seulement 1500 mots.

#### 2.2.2.1 Formalismes utilisés dans la définition DEC

Une définition DEC forme un seul bloc de texte. Toutefois, on utilise des procédés typographiques pour distinguer certaines composantes de la définition : les composantes présupposées, les composantes optionnelles, les composantes faibles et les composantes exprimant un lien de polysémie. Nous décrirons chacune de ces composantes et en donnerons des exemples. Les composantes qui nous intéressent sont soulignées dans les définitions.

Un **présupposé** est une composante du sens d'une phrase ou d'une lexie qui n'est pas niée lorsque cette phrase ou cette lexie est niée. Les présupposés seront présentés en détail dans la section 3.2. Dans les définitions DEC des DECFC, les deux barres « || » isolent la partie présupposée à la gauche du reste de la définition. Cet élément de formalisme est désormais

remplacé par les signes de ponctuation suivants : « [...] », qui encadrent la partie présupposée. Voici un exemple de présupposé :

SE BERCIER<sup>1.1</sup> (DECFC-IV : 118)

[*Grand-mère se berce devant le foyer.*]

*X se berce dans Y* = [ Se trouvant sur un siège Y et balançant Y, ] | personne X cause de se balancer doucement et régulièrement dans le but de donner à X une sensation agréable.

**Définition 11 : Définition DEC de SE BERCIER<sup>1.1</sup>**

Soit la phrase :

(3) *Quand elle parle, grand-mère ne se berce pas.*

Même si SE BERCIER est nié, on imagine que grand-mère se trouve sur un siège qui permettrait de se bercer.

Les composantes optionnelles font partie du sens de la lexie définie à moins que l'on ne mentionne explicitement le contraire dans le contexte plus ou moins rapproché de l'utilisation de la lexie dans le discours. Les parenthèses « (...) » identifient les composantes optionnelles, en voici un exemple :

BOUILLOIRE<sup>a</sup> (DECFC-IV : 140)

[*J'ai mis la bouilloire sur le feu.*]

*Bouilloire de X [pour Y]* = Ustensile<sup>1</sup> de cuisine destiné à ce qu'une personne X y fasse bouillir<sup>1.1a</sup> de l'eau Y [...] – récipient (métallique) fermé ayant un bec et une anse supérieure.

**Définition 12 : Définition DEC de BOUILLOIRE<sup>a</sup>**

Ainsi, une bouilloire est, par défaut, faite de métal, mais il existe des bouilloires qui ont un revêtement de plastique et on pourrait imaginer une bouilloire faite d'un autre matériau. Si l'on parle d'une *bouilloire de plastique*, il s'agit bien de la lexie BOUILLOIRE<sup>a</sup>, pour laquelle la composante 'métallique' est ignorée, comme en témoigne le test de Green-Apresjan<sup>5</sup> qui prévoit que, si pour une lexie on peut utiliser deux occurrences coordonnées sans faire de jeu de mots, alors on est en présence d'une seule lexie :

<sup>5</sup> Ce test est une reformulation d'un critère proposé par Apresjan (1974) et dont le prototype se trouvait dans Green (1969).

(4) *Jean a apporté deux bouilloires, une de métal et une de plastique.*

La **composante faible** est une composante qui peut s’effacer dans des conditions précises et spécifiées dans la définition. Elle se trouve entre parenthèses en caractères gras, avec un indice qui spécifie les conditions qui l’annulent. Voici un exemple de composante faible :

ÉTUDIANT

*étudiant en Y de Z* = ‘personne (de sexe masculin)<sub>gén</sub> qui étudie la matière Y dans un établissement d’enseignement supérieur Z’

**Définition 13 : Définition DEC de ÉTUDIANT**

La composante ‘de sexe masculin’ peut s’estomper si on fait un usage générique ou pluriel de la lexie ÉTUDIANT, comme dans la phrase :

(5) *Les étudiants vont à la cafétéria.*

Il peut s’agir d’étudiants des deux sexes.

Avant de voir les composantes exprimant un lien de polysémie, présentons les deux types de liens de polysémie dont il est question dans l’ILEC : la **métonymie** et la **métaphore**.

La composante sémantique partagée entre L1 et L2 qui fait qu’il y a polysémie au sein d’un vocable est appelée **pont sémantique**.

**Métonymie** : Le lien entre la lexie L1 et la lexie L2, qui partagent un lien métonymique, est basé sur une contiguïté de leurs référents (contenant~contenu, tout~partie, etc.). Les liens métonymiques ne sont pas marqués de façon explicite dans le DEC.

**Métaphore** : Le lien entre la lexie L1 et la lexie L2, qui partagent un lien métaphorique, est basé sur une ressemblance entre leurs référents. Dans le DEC, les composantes qui indiquent la présence d’un lien métaphorique se trouvent entre crochets « [...] » et sont introduites par le sémantème ‘comme si’. Voici un exemple de lien métaphorique :

BERCEAU**II.1** (DECFC-IV : 114)

[*La Nouvelle-Orléans, berceau du blues, est en deuil*]

**II.1.** *Berceau* [X] de Y = Lieu X où a commencé l'existence de Y [comme si X était le berceau**I** de Y].

**Définition 14 : Définition DEC de BERCEAU**II.1****

La composante ‘comme si X était le berceau**I** de Y’ exprime le lien métaphorique qui existe entre les lexies BERCEAU**I** [*Le bébé est couché dans le berceau*] et BERCEAU**II.1**.

Le lien métaphorique est différent des présupposés, des composantes faibles et des composantes optionnelles. Ce lien existe dans le lexique mental du locuteur d'une langue, il rend explicites les ponts sémantiques qui permettent de faire des jeux de mots et des plaisanteries. Aussi, il fait partie de la définition d'une lexie, mais ne se trouve pas dans la SSém d'une phrase dans laquelle cette lexie est employée. Il en est de même pour tous les liens de polysémie, nous référons le lecteur à Barque (2008) pour plus de détails sur les différents liens de polysémie.

### 2.3 *La BDéf*

Nous voici à la présentation de la troisième formalisation étudiée. Il s'agit de l'une des plus récentes formalisations développées dans la TST.

#### 2.3.1 **Caractérisation de la définition BDéf**

##### 1) À qui s'adresse-t-elle ?

Le format de définition BDéf a été créé comme outil lexicologique pour le traitement automatique du langage, c'est-à-dire qu'un programme informatique peut interpréter son contenu et calculer des phénomènes sémantiques de façon automatique. Toutefois, le format BDéf vise à ce que le contenu demeure suffisamment accessible pour qu'un lexicographe puisse lui aussi l'interpréter et l'éditer.

##### 2) À quel type d'entité linguistique s'applique-t-elle ?

À notre connaissance, le format BDéf n'a jamais été utilisé pour formaliser le sens de phrases. En théorie, cela serait tout à fait possible, mais il faudrait en remanier le

format, renommer certains de ses blocs et se doter de nouveaux outils, nécessaires, entre autres, pour pouvoir traiter toutes les oppositions communicatives.

### 2.3.2 Description de la définition BDéf

Altman & Polguère (2003) proposent une autre formalisation dans la BDéf, une base de données de définitions hautement formalisées. La BDéf fait aussi l'objet des travaux de Barque (2003, 2008), Barque & Nasr (2005) et de Barque & Polguère (2004).

Dans la BDéf, on reprend les définitions du DECFC, que l'on réécrit dans un langage encore plus contrôlé, duquel la flexion, les pronoms et les synonymes sont complètement évacués et dans lequel les locutions sont explicitées (elles sont encadrées par des traits de soulignement, par ex. \_porter\_atteinte\_).

Le format utilisé dans la BDéf constitue en quelque sorte un format intermédiaire entre la définition lexicographique et le réseau sémantique. Comme le réseau sémantique, il permet de voir clairement les sémantèmes impliqués dans la définition du lexème; par contre, comme la définition verbale, il possède l'avantage de la linéarité, qui impose un ordre de lecture. Aussi, ce format est plus près d'une définition verbale que celui du réseau sémantique, et de ce fait, il satisfait mieux l'intuition linguistique du chercheur.

Outre la section « Définition » décrite ci-après, un article BDéf contient le nom de la lexie et son numéro lexicographique distinctif, sa caractérisation grammaticale, son étiquette sémantique, sa forme propositionnelle, ainsi que des exemples d'utilisation.

#### 2.3.2.1 La section Définition

L'une des particularités de la BDéf, qui la démarque par rapport aux formalisations traditionnelles de la LEC, est la division d'une définition en blocs. La définition BDéf comporte quatre blocs définitionnels :

- 1) le bloc central ;
- 2) le bloc des différences spécifiques [= nos composantes périphériques] ;
- 3) le bloc de typage d'actants ;
- 4) le bloc de relations sémantiques entre les actants.

Il est intéressant de remarquer que les concepteurs de la BDéf ont choisi de séparer le typage des actants du corps de la définition plutôt que d'intégrer ces données en marquant les restrictions sur les actants à l'aide d'une convention typographique. Aussi, il est à noter que les représentations linéaires ne favorisent pas l'explicitation des liens qui existent entre les différentes composantes d'une définition. C'est le cas notamment pour les relations entre les actants. Il s'agit là d'un concept lexicographique qui est mieux véhiculé par une représentation multidimensionnelle telle que le réseau sémantique. Dans la BDéf, on dédie un bloc aux relations entre les actants pour être en mesure d'explicitier ces liens.

Voici un extrait de la fiche de la lexie OURAGANI. Dans cet extrait, les quatre blocs de la section Définition, présentés ci-dessus sont identifiés :

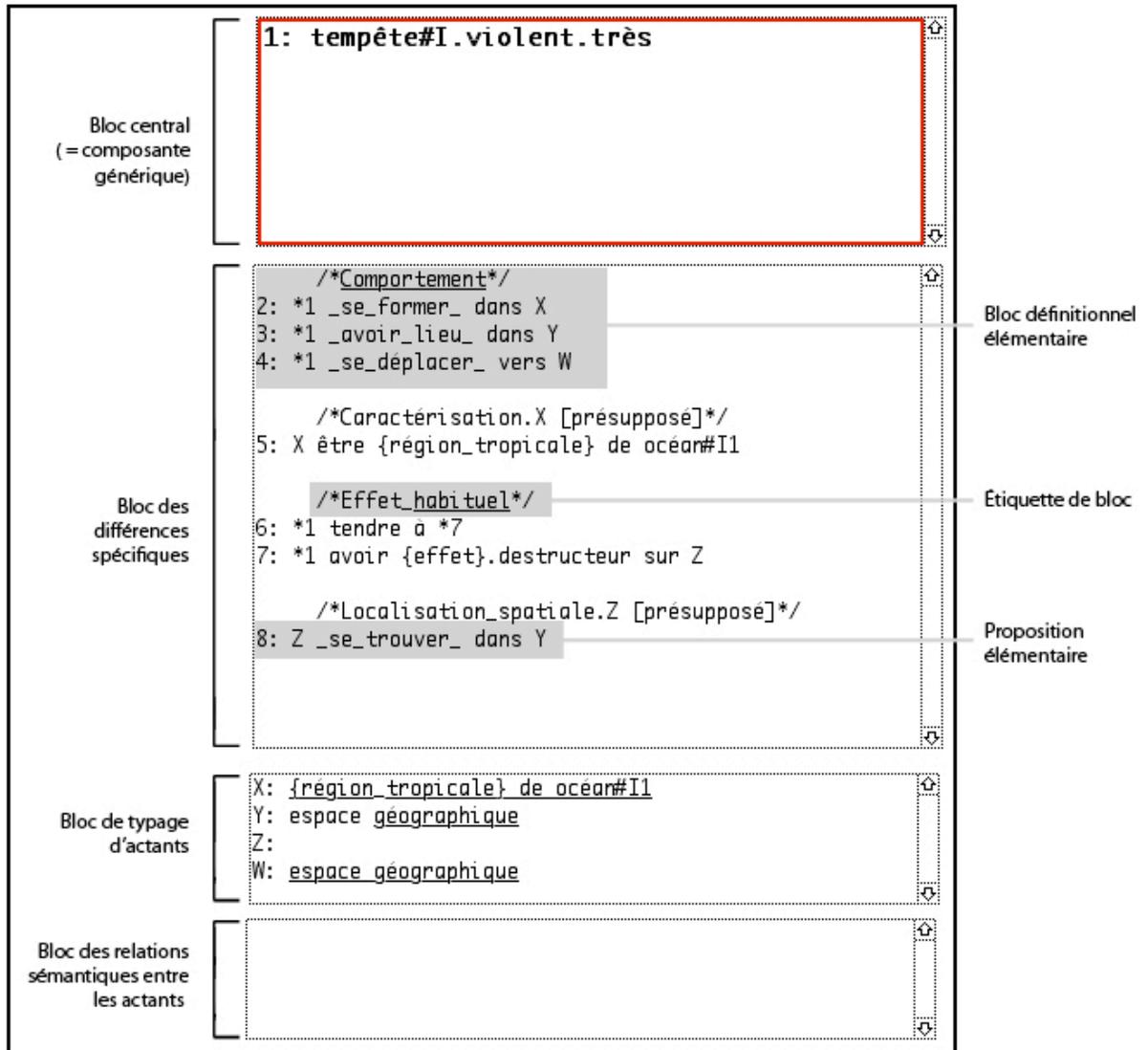


Figure 9 : Présentation des différentes composantes d'une définition BDéf

Le premier bloc, le bloc central, correspond à la composante générique du sens lexical défini. Le deuxième bloc est celui des différences spécifiques (composantes périphériques). Celui-ci se divise en blocs définitionnels élémentaires, séparés par un retour de chariot. Ces blocs sont formés de propositions élémentaires, qui occupent chacune une nouvelle ligne. Ces propositions sont elles-mêmes composées de sémantèmes et sont regroupées selon leur rôle informationnel, c'est-à-dire selon le type de spécificité qu'elles apportent à la composante générique. Il peut s'agir de la description des différentes parties de l'objet à

définir, de sa localisation, de son but, etc. La définition d'un fait peut inclure une cause, un but, une localisation, etc. Le rôle informationnel du bloc est identifié par ce qui est appelé une *étiquette de bloc* dans la terminologie de la BDéf, nous parlerons plutôt d'*en-tête de bloc* pour éviter toute confusion avec les étiquettes sémantiques, un type de composantes lexicologiques que nous introduirons à la section 4 du chapitre 2. Les en-têtes de bloc sont notés de la façon suivante : /\*En-tête de bloc\*/. Dans l'exemple de la Figure 9 ci-dessus, les en-têtes de blocs sont /\*Comportement\*/, /\*Caractérisation\*/, etc.

Nous avons maintenant vu les trois formalisations qui font l'objet de ce travail. Avant de pouvoir aborder les différences entre les trois formalisations de RSém décrites dans cette section, nous introduirons quelques notions qui sont nécessaires à notre exposé.

### 3 La structure communicative (SSém-Comm)

Dans la TST, on postule la structure communicative (SSém-Comm) pour rendre compte des phénomènes communicatifs qui se manifestent au sein des énoncés. Ainsi, *Communicative organisation in natural language* (Mel'čuk, 2001), dorénavant CONL, un ouvrage dédié à la description de la structure communicative, s'attarde très peu sur le rôle et la représentation de la SSém-Comm dans les RSém de lexies. En fait, dans la description des oppositions communicatives, aucun exemple ne concerne les RSém de lexies.

Tout de même, trois des oppositions communicatives s'appliquent aussi aux RSém de lexies. Ainsi, la présupposition et la locutionnalité s'appliquent surtout aux RSém de lexies, mais aussi au RSém de phrases. La thématicité est pertinente pour les deux types de RSém (lexies et phrases).

Pour la suite de notre exposé, une brève présentation de ces trois oppositions s'impose. Pour plus de détails sur ces oppositions communicatives ou sur la structure communicative en général, nous renvoyons le lecteur à CONL (*Id.*), cité plus haut.

#### 3.1 La thématicité

La thématicité a trois valeurs : le thème (**T**-Sém), le rhème (**R**-Sém) et les spécifieurs. Voici les définitions des deux premiers concepts, données dans CONL :

Rhème sémantique :

Le *rhème sémantique* d'une phrase **P** est la partie 'σ' de sa SSém 'S' que le locuteur présente comme étant COMMUNIQUÉE par **P** (Mel'čuk, 2001 : 102).

Thème sémantique :

Le *thème sémantique* d'une phrase **P** est la partie 'σ' de sa SSém 'S' que le locuteur présente comme étant ce À PROPOS DE QUOI LA SEM-R('S') EST COMMUNIQUÉE (*Ibid.* : 104).

Le rhème possède, en règle général, un **focus rhématique**. Le focus rhématique est la partie du sens d'un énoncé (plus précisément du Rhème sémantique) qui répond à une question sous-jacente, c'est-à-dire une question à laquelle cet énoncé peut être une réponse convenable. Polguère (1990 : 45) nous donne l'exemple suivant, les questions sous-jacentes possibles sont précédées par une flèche :

(6) *Les plans de la fusée pour la Lune sont introuvables.*

← 'Qu'est-ce qui est arrivé aux plans de la fusée pour la Lune?'

focus rhématique : *introuvables*

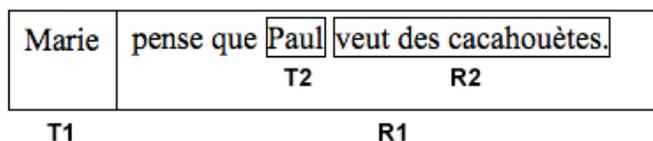
← 'Qu'est-ce qui est introuvable?'

focus rhématique : *les plans*

etc.

La division thème-rhème peut être récursive. Cela signifie que le rhème ou le thème d'un énoncé peut lui-même contenir un rhème et un thème. Par exemple, dans cette phrase complexe avec subordonnée :

(7)



Puisqu'un énoncé communique toujours quelque chose, le rhème y est obligatoire, mais pas le thème. Mel'čuk (2001 : 111-112) mentionne deux types d'énoncés sans thème (appelés

communicativement inarticulés). Il y a d'abord les énoncés communicativement inarticulés indépendants du contexte, aussi appelés phrases rhématiques. Ces énoncés pourraient répondre à des questions du type : *Que s'est-il passé? Quoi de neuf?* Ils se trouvent souvent au tout début du discours. Mel'čuk (*Ibid.*) donne pour exemple les phrases suivantes :

(8) *Il pleut.*

(9) *Beaucoup de phénomènes sémantiques ne sont pas couverts dans ce livre.*

On y mentionne aussi les phrases du type : « *Il était une fois...* ». Toutes ces phrases ne sont composées que d'un rhème.

Il existe également des énoncés communicativement inarticulés dépendants du contexte. Voici quelques-uns des exemples traduits, tirés de Mel'čuk (*Ibid.*), le contexte est entre crochets :

(10) [*Désolé, je dois partir.*] *Le taxi attend.*

(11) [*– Pourquoi es-tu si triste?*] – *John est mort hier.*

Les phrases (10) et (11) ne contiennent pas de thème.

Les spécifieurs, qui constituent la troisième valeur de la thémativité, sont de trois types. Les **circonstanciels** indiquent différentes circonstances dans lesquelles se déroule un fait ; ils correspondent aux compléments circonstanciels en grammaire traditionnelle (par ex. *Demain, ...* ou encore *À New York, ...*), les **caractériseurs** marquent l'attitude subjective du locuteur vis-à-vis de son énoncé et/ou du contenu de ce dernier (*Malheureusement..., Si je peux me permettre...*) et les **connecteurs** marquent un lien logique ou rhétorique entre deux énoncés (*car, puisque, bref, de plus*). Seuls les connecteurs sont pertinents pour la description de la SSém-Comm de lexies. Voici un exemple dans lequel chacune des valeurs de la thémativité est explicitée :

(12) Hier midi<sub>Spécifieur-Sém</sub>, Jean<sub>Thème-Sém</sub> a mangé des cacahouètes<sub>Rhème-Sém</sub>.

← 'Qu'est-ce que Jean a fait?'

Dans CONL, le formalisme utilisé pour marquer la thémativité est l'encadré étiqueté, et ce, autant dans le réseau que dans les exemples linguistiques. On pourrait employer le même formalisme dans la RSém d'une lexie.

### 3.2 La présupposition

La présupposition, en tant qu'opposition communicative, comprend deux valeurs : le **posé** et le **présupposé**. Le posé est la valeur par défaut et seule la valeur présupposé nous intéresse ici. Comme nous l'avons déjà dit, le présupposé est la partie du sens d'un énoncé ou d'une lexie qui ne peut être niée lorsque l'énoncé/la lexie l'est. Par exemple, la lexie **SE CASSER LE COU**, dont voici la définition DEC, contient une composante présupposée :

**SE CASSER LE COU** (DECFC-IV : 171)

*X se casse le cou en Y-ant* = ||En faisant Y, || personne X se heurte à un obstacle de façon brutale, le choc causant un dommage corporel assez grave à X.

**Définition 15 : Définition DEC de SE CASSER LE COU**

Soit la phrase :

(13) *Georges s'est cassé le cou en ski alpin.*

Si on nie cette phrase, on obtient :

(14) *Georges ne s'est pas cassé le cou en ski alpin.*

La composante 'Georges faisait du ski alpin' n'est pas niée.

Dans CONL, le formalisme utilisé dans un réseau sémantique pour identifier une composante présupposée est une bulle ombragée. Dans les exemples, on met la partie présupposée entre crochets comme ceux-ci : ||présupposition||.

### 3.3 La locutionnalité

La locutionnalité a trois valeurs : **communiqué**, **signalé** et **performatif**. Il faut tenir compte du fait que la locutionnalité, contrairement aux autres oppositions communicatives, s'applique à l'ensemble de la SSém d'une lexie et non à un sous-ensemble de celle-ci (sous-réseau). Ainsi, dans une définition DEC, on exprime la valeur de la locutionnalité au niveau de la lexie, c'est-à-dire qu'elle vaut pour l'ensemble de la SSém de la lexie et non seulement pour une composante sémantique de la lexie. La locutionnalité d'une lexie ou

d'une phrase est exprimée dans le **cadre illocutoire** de la SSém de cette dernière, c'est-à-dire dans un « en-tête » pour la définition, qui annonce sa valeur illocutoire.

Un élément communiqué contient la composante sémantique 'je veux que tu saches que...'. Il s'agit de la valeur non marquée de la locutionnalité. Un élément signalé ne contient pas cette composante, il ne vise pas à communiquer quelque chose à quelqu'un. Par exemple, *Ça me dégoûte!* est communiqué, son cadre illocutoire est 'Je veux que tu saches que...' tandis que *Beurk!*, qui a le même sens, est signalé. Son cadre illocutoire est 'Je signale que...'. Par contre, les **lexies descriptives** (dont le signifié est communiqué) n'ont pas besoin d'être marquées comme telles, puisqu'il s'agit de la valeur de locutionnalité par défaut.

Un performatif est une lexie telle que celui qui la prononce à la première personne et au présent de l'indicatif fait l'action qu'elle décrit en la prononçant : *je jure, je promets, je vous déclare mari et femme*, etc.

Une lexie dont le signifié est signalé est un **signalitif**. Les performatifs sont eux aussi des **lexies non descriptives** et leur signifié devrait contenir une indication à cet effet.

#### **4 Les étiquettes sémantiques**

Nous introduisons maintenant un type de composantes sémantiques qui sera utilisé dans nos formalisations : les étiquettes sémantiques. Une étiquette sémantique (Milićević, 1997; Polguère, 2003b) est une formule qui permet de référer au sens d'une lexie sans faire appel à la définition complète de celle-ci. Elle correspond à la composante générique du sens d'une lexie. Les étiquettes sémantiques ont, entre autres, une fonction classificatrice fort utile dans les applications de TALN. Elles sont utilisées dans le DiCo, le DiCE (Alonso Ramos, 2003; Wanner *et al.*, 2004) (un dictionnaire de collocations de l'espagnol de type explicatif et combinatoire), la BDéf, le *Dire Autrement* et le LAF. Dans le DiCo, elles font office de définitions minimales puisqu'elles constituent la seule information sémantique donnée pour les lexies traitées. Voici quelques exemples d'étiquettes utilisées dans le DiCo pour caractériser de façon approximative le sens des lexies (elles se trouvent dans la colonne de droite) :

lexie vocalbe	lexie num	lexie formuleEtiquette
À TOUT CASSER(1)	I.1	haut degré
ABAISSEMENT	I.1	processus
ABANDON	I.1	arrêt volontaire
ABÎME	II.1	situation difficile symbolisée par un élément du relief
ACCUSATION	I.1	énoncé
ACCUSER	I.1	dire qqch.
AMI	I.1	individu qui a une certaine relation interpersonnelle avec un autre

Figure 10 : Extrait du DiCo montrant les colonnes "vocalbe", "numéro de vocalbe" et "étiquette"

Les étiquettes sémantiques s'organisent en une hiérarchie (Polguère, 2009) (<http://olst.ling.umontreal.ca/dicopop-laf/hierarchie.php>). Pour être reconnu comme une étiquette sémantique, un sens lexical doit correspondre à la composante générique d'au moins deux lexies dans une langue *L*. Voici un extrait de la hiérarchie :

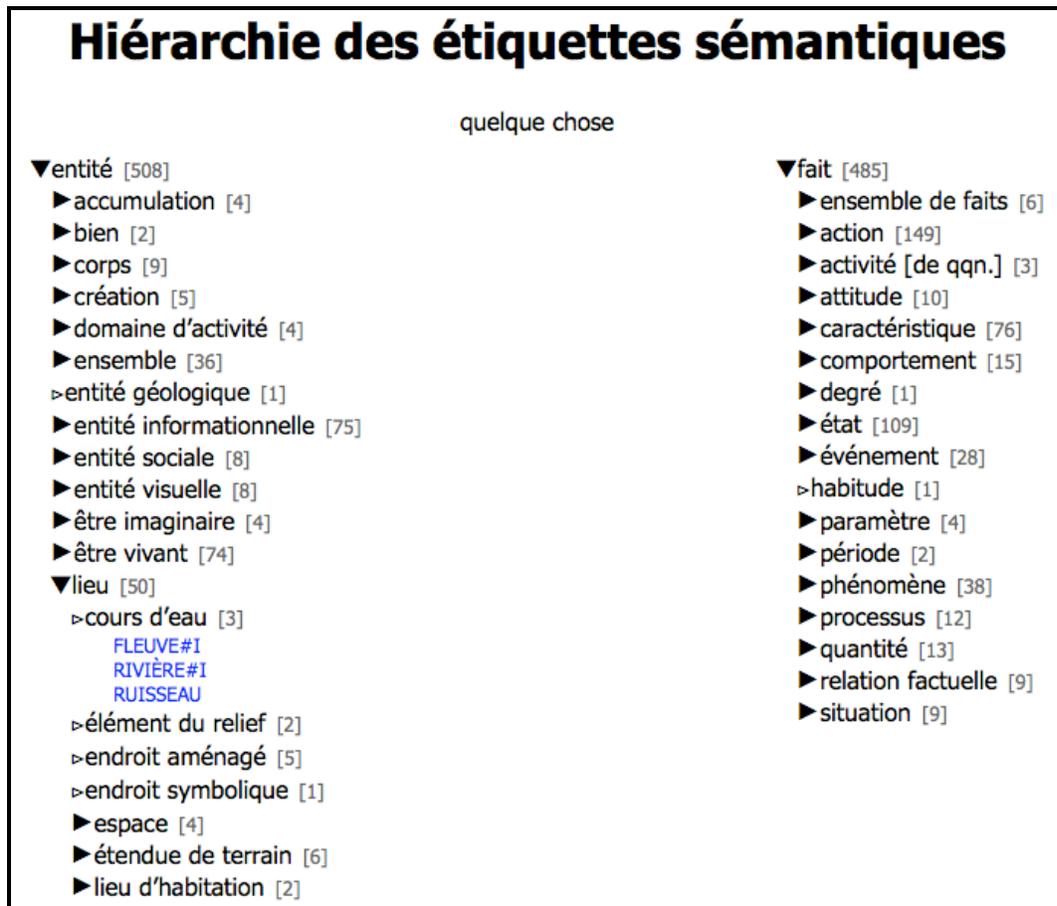


Figure 11 : Extrait de la hiérarchie d'étiquettes sémantiques

Au sommet de la hiérarchie se trouve l'étiquette sémantique *quelque chose*. De là, la hiérarchie se sépare en deux branches principales : *entité* et *fait*. Nous voyons que sous l'étiquette sémantique *lieu* se trouve l'étiquette *cours d'eau*, sous laquelle sont représentées en petites majuscules les lexies dont le sens a été décrit au moyen de cette étiquette sémantique (FLEUVE#I, RIVIÈRE#I, RUISSEAU).

La hiérarchie d'étiquettes a plusieurs fonctions, dont celles de permettre une structuration du lexique et le calcul de l'héritage de propriétés lexicales. Elle permet également une normalisation de la description lexicographique.

Dans le cadre de ce travail, nous utilisons les étiquettes sémantiques en tant que composantes qui contraignent les valeurs que peut prendre un actant, nous reviendrons sur les contraintes sur les actants et le typage d'actants au chapitre 2, à la section 5.4. Aussi, la hiérarchie d'étiquettes sémantiques permet d'illustrer certains phénomènes qui concernent

les composantes génériques, comme les composantes génériques doubles, représentées par des conjonctions et des disjonctions d'étiquettes (voir plus loin, à la section 5.5 du chapitre 2).

## 5 Questions théoriques non résolues

### 5.1 *La représentation de la thémativité dans une RSém de lexie*

Dans la TST, il n'y a pas de travaux poussés concernant le rôle de la thémativité dans la RSém d'une lexie. Aussi, pour la définition BDéf ou la définition DEC, on ne propose aucun formalisme pour expliciter la thémativité. Le rôle que joue la thémativité dans l'organisation du sens d'une lexie n'a pas été déterminé, la façon ou les critères pour le représenter non plus. Pourtant, l'opposition thémativité semble être pertinente pour la RSém d'une lexie, et ce, pour toutes les parties du discours. Il apparaît que la représentation de la thémativité est surtout intéressante dans une perspective interlinguistique, dans laquelle on voudrait expliciter la structure d'une langue par rapport à la structure d'une autre. Comme il s'agit d'un problème de taille, nous ne proposerons pas ici de formalisme pour représenter la thémativité, mais soulignons le fait que cette problématique devrait être étudiée plus en profondeur.

### 5.2 *Le rôle de la chronologie des faits dans l'ordre de lecture des composantes*

« Quelles sont les relations temporelles entre les faits mis en jeu par le sémantisme de L? » (Mel'čuk *et al.*, 1995 : 110). On pose cette question dans l'ILEC, mais on ne propose pas de formalisme à utiliser pour expliciter ces relations temporelles.

Par exemple, dans la définition BDéf de OURAGAN<sub>I</sub>, il y a un élément chronologique que l'on tente de respecter dans l'ordre et dans l'organisation du bloc /\*Comportement\*/.

```

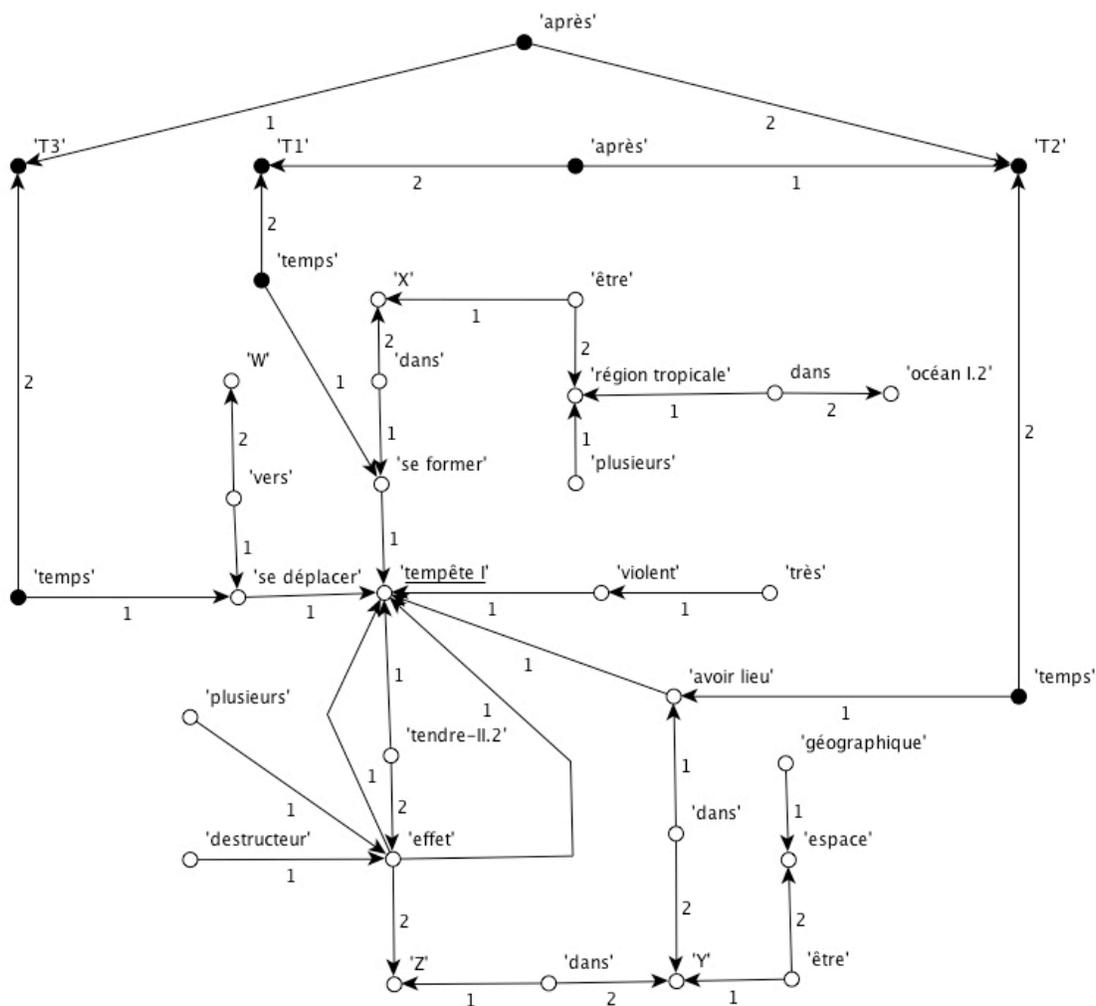
/*Comportement*/
2: *1 _se_former_ dans X
3: *1 _avoir_lieu_ dans Y
4: *1 _se_déplacer_ vers W

```

Figure 12 : Le bloc /\*Comportement\*/ dans la définition BDéf de OURAGAN<sub>I</sub>

Si ce n'était pas de cette suite chronologique, la composante '\*1 \_avoir\_lieu\_ dans Y' se trouverait peut-être dans un bloc à part, portant l'en-tête /\*Localisation\_spatiale\*/. Cette notion d'ordre chronologique est très intéressante dans le cadre de nos questionnements sur l'ordre de lecture des composantes de la SSém.

Il faudrait ajouter des composantes temporelles dans les formalisations de OURAGANI. En fait, ces composantes ne constituent rien de nouveau dans la formalisation sous forme de réseau, mais leur impact sur l'orientation de la lecture du réseau n'a pas été déterminé :



**Réseau sémantique 3 : Réseau sémantique de la lexie OURAGANI avec indications temporelles**

Dans le Réseau sémantique 3, les nœuds qui indiquent la temporalité ont été remplis en noir afin de les mettre en évidence. Il n'existe pas de façon d'expliciter la temporalité dans les définitions DEC et les définitions BDéf. Du moins, on ne semble pas les utiliser dans les

DECFC publiés. Il faudrait étudier le sujet plus en profondeur pour être en mesure d'expliquer comment fonctionne la temporalité dans l'ordonnement des sémantèmes d'une définition.

### 5.3 *Les conjonctions*

Les trois composantes du bloc élémentaire /\*Comportement\*/ dans la définition BDéf de OURAGAN<sup>I</sup>, montrées à la Figure 12, sont liées par un 'et' dans la définition du DEC. Ce 'et' est pourtant vide de sens, il ne sert qu'à rendre la lecture plus naturelle. Il n'a pas d'équivalent ni dans le réseau sémantique, ni dans la définition BDéf. En théorie, on pourrait l'omettre.

D'autre part, il faudrait expliciter les « vrais » connecteurs logiques, ainsi que la façon dont ils interagissent avec la négation. Par exemple, dans l'article de Dostie *et al.* (1999), on mentionne la différence entre 'et' et 'et aussi'. Ainsi, le connecteur 'et' correspondrait à plusieurs égards à la conjonction logique. Par contre, elle n'est pas tout à fait équivalente à la conjonction logique puisque la négation d'un énoncé contenant une conjonction logique (*A et B.*) permet les trois interprétations suivantes :

'Non-A et Non-B'

'Non-A et B'

'A et Non-B'

Or, pour certaines définitions contenant une conjonction, on ne veut pas permettre ces trois interprétations. Parfois, une des propositions en particulier doit être niée pour que le sens de la lexie soit nié. Par exemple, la lexie REPROCHER<sup>1a</sup> :

(15) *Marie reproche à Pierre son retard.* ≈ 'Marie a présent dans sa conscience qu'elle n'excuse<sup>2</sup> pas Pierre pour son retard, **et** Marie croit qu'il doit être conscient du caractère relativement mauvais de son retard'

Si on nie (15), on obtient :

(16) *Marie ne reproche pas à Pierre son retard.*

La phrase (16) peut être analysée des deux façons suivantes :

1) Marie **n'a pas** présent dans sa conscience qu'elle n'excuse<sup>2</sup> pas Pierre pour son retard, **et** Marie croit qu'il doit être conscient du caractère relativement mauvais de son retard'

(‘Non-A et B’)

2) Marie **n’a pas** présent dans sa conscience qu’elle n’excuse<sup>2</sup> pas Pierre pour son retard, **et** Marie **ne** croit **pas** qu’il doit être conscient du caractère relativement mauvais de son retard’  
(‘Non-A et Non-B’)

Mais (16) ne peut pas être interprétée comme étant :

3) Marie a présent dans sa conscience qu’elle n’excuse<sup>2</sup> pas Pierre pour son retard, **et** Marie **ne** croit **pas** qu’il doit être conscient du caractère relativement mauvais de son retard’  
(‘A et Non-B’)

Ainsi, dans la définition d’une lexie comme REPROCHER<sup>1a</sup>, on préfère utiliser un connecteur tel que ‘et aussi’ pour souligner le caractère asymétrique de la conjonction. Dans Dostie *et al.* (*Ibid.*), on mentionne aussi l’existence des connecteurs ‘et après’ et ‘et par conséquent’. La construction d’une typologie de ces connecteurs est nécessaire. Nous ne la ferons pas dans le cadre de ce travail.

#### 5.4 Caractérisation de variables vs typage de variables

Comme nous l’avons vu, une définition BDéf contient un Bloc de typage des actants. Ce bloc comprend chacune des variables actanciennes de la lexie décrite auxquelles sont associées des contraintes taxinomiques limitant les valeurs que peut prendre l’instanciation de cette variable. On parle aussi d’*étiquetage sémantique* puisque les contraintes associées aux variables correspondent à des étiquettes sémantiques ou, dans certains cas, à une disjonction ou conjonction d’étiquettes sémantiques.

Par exemple, le typage des actants de la lexie PEIGNE<sup>1a</sup> [*X peigne Y de Z*] est représenté dans le bloc de typage suivant :

X: individu	
Y:	
Z: <u>être humain ou animal</u>	

Figure 13 : Le bloc de typage d’actants dans la définition BDéf de PEIGNE<sup>1a</sup>

Dans l’exemple ci-dessus, individu, être humain et animal correspondent chacun à des étiquettes sémantiques.

Toujours dans la BDéf, dans le Bloc des différences spécifiques, on trouve des blocs définitionnels élémentaires dont l'en-tête est */\*Caractérisation.Variable\*/*, comme dans la définition BDéf de **OURAGANI** :

```
/*Caractérisation.X [présupposé]*/
5: X être {région_tropicale} de océan#I1
```

**Figure 14 : Bloc élémentaire */\*Caractérisation\*/* dans la définition BDéf de **OURAGANI****

Ce bloc signifie que l'une des composantes périphériques de **OURAGANI** est la caractérisation sémantique de la position actancielle 'X' spécifiant que 'X est une région tropicale d'un océan**I1**'. Les blocs définitionnels de ce type contiennent de l'information concernant une variable et cette information est considérée comme étant de nature différente de celle renfermée dans le Bloc de typage des actants. Il semble que le typage des actants ne fasse pas vraiment partie du sens de la lexie, il s'agit seulement d'une indication quant à son usage. La caractérisation d'actant, pour sa part, fait pleinement partie du sens de la lexie et constitue une composante périphérique dans la définition.

Il n'est pas toujours facile de distinguer les deux. Tout de même, la distinction entre typage d'actant et caractérisation d'actant devrait être introduite dans les autres formalisations. Il faut d'abord définir des critères précis pour distinguer d'une part les caractérisations d'actants, qui se trouvent dans le corps de la définition, et d'autre part les typages d'actants, qui sont une information supplémentaire permettant de restreindre le choix des éléments pouvant remplir la position actancielle. Cela est essentiel puisque les caractérisations constituent une composante à part entière de la définition, tandis que les contraintes actanciennes, qui correspondent à des étiquettes sémantiques, n'ont pas le même statut que les autres sémantèmes de la définition.

### **5.5 Les lexies à composante générique double**

Les définitions DEC de plusieurs artefacts mettent en jeu un formalisme particulier, comme nous l'avons vu dans la définition de **PEIGNE1a**. Sa composante générique est 'ustensile**1**', puisqu'on met la fonction de l'artefact ('ustensile', 'instrument de cuisine') comme composante générique, mais 'objet' est aussi instancié pour permettre d'énoncer des propriétés qui se rapportent à son état d'objet plutôt que d'ustensile. Les deux sémantèmes sont reliés par le tiret long « – », qui représente le sémantème 'être', un usage qui pourrait être critiqué puisqu'un symbole ne devrait pas remplacer un sens. Voici de nouveau la

définition de PEIGNE**1a** :

PEIGNE**1a** :

*Peigne de X [pour peigner Y de Z] = Ustensile**1** de toilette destiné à ce que l'individu X peigne**1** les cheveux ou le poil Y de Z – objet plat assez rigide censé être tenu à la main**1.a** par X, et muni d'une ou de deux rangée(s) de dents**II.1**.*

Deux analyses sont possibles. On peut penser que 'objet' se trouve en fait dans la décomposition sémantique de 'ustensile', mais qu'il doit en être extrait pour permettre de décrire la forme de l'ustensile en question. On pourrait aussi penser qu'il s'agit d'une deuxième composante générique, puisque c'est 'objet' qui détermine la combinatoire libre de la lexie. Pour la représentation sémantique, il est essentiel de définir ce phénomène davantage puisque nous devons établir la façon de représenter plusieurs composantes génériques dans une formalisation.

Les lexies à composante générique double sont documentées dans Polguère (2003b), Barque (2008) et Milićević & Polguère (2010). Polguère (*Ibid.* : 58-59) développe le sujet des lexies à composante générique multiple et de l'implication que cela a sur leur étiquette sémantique. Il distingue deux types d'étiquettes sémantiques de lexies : les **étiquettes avec disjonction exclusive** et les **étiquettes avec disjonction inclusive**. Pustejovski (1995 : 92-95), dans le cadre de la théorie du Lexique génératif, note ce phénomène et le représente à l'aide des types pointés (*dotted types*).

Polguère donne comme exemple, pour les étiquettes avec disjonction exclusive, tous les noms de profession (DOCTEUR, AVOCAT, NOTAIRE). Ainsi, les noms de profession peuvent désigner soit la personne, soit la profession, ils posséderont donc deux composantes génériques. Par contre, seulement une des significations peut être employée à la fois. On les regroupe tout de même dans une même lexie pour éviter la redondance. Dans ce cas, la composante générique est 'personne ou profession', avec un 'ou' exclusif.

Ainsi, la lexie PEIGNE**1a**, dans nos exemples, peut être utilisée pour ses deux étiquettes (ses deux composantes génériques) à la fois : 'ustensile' et 'objet'. Nous croyons qu'il ne s'agit

pas d'une disjonction puisqu'il y a un lien particulier en les deux composantes génériques; la composante générique de PEIGNE**1a** sera plutôt 'ustensile est un objet'.

Nos formalismes devront expliciter la différence entre les lexies qui ont des étiquettes à disjonction exclusive ou inclusive et les étiquettes du type de celle de PEIGNE**1a**. Par contre, l'ensemble des cas de figure ne sera pas traité dans le cadre de ce travail (conjonction d'étiquettes et autres types d'étiquettes).

## CHAPITRE 3 Les disparités dans les formalisations traitées

### 1 Introduction

Revenons aux trois formalisations qui nous intéressent. Il existe plusieurs différences dans la nature de l'information véhiculée par chacune de celles-ci. Bien sûr, certaines formalisations se prêtent mieux à la mise en évidence de certains phénomènes. Aussi, certaines de ces formalisations ne permettent pas d'exprimer de façon efficace et élégante tout ce que l'on pourrait vouloir y exprimer. Par contre, ces différences ne sont pas toujours justifiées par des limites inhérentes à la formalisation elle-même, ni par des choix ou des objectifs théoriques : ce sont les cas qui nous intéressent.

Afin d'illustrer nos propos, voici trois exemples sur lesquels nous reviendrons en détail dans les sections à venir.

- 1) Dans la définition DEC, aucun élément de formalisme ne fait ressortir ce qui constitue la composante générique ou les composantes périphériques.
- 2) Dans le réseau sémantique, on souligne le NDg. Celui-ci n'est identifié ni dans la définition DEC, ni dans la BDéf.
- 3) Dans les définitions BDéf, on isole la composante générique, qui, rappelons-le, ne correspond pas au NDg. La composante générique n'est pas identifiée dans les deux autres formalisations.

Il est de notre avis que chacune des formalisations du sens d'une lexie L devrait présenter toutes les informations nécessaires à l'utilisation de L et que les mêmes informations devraient être données dans chacune des formalisations, lorsque rien ne s'y oppose. Par exemple, le réseau, la définition DEC et la définition BDéf pourraient identifier le NDg et l'identification du NDg pourrait et devrait se faire de la même façon d'une formalisation à une autre; nous utiliserons le soulignement comme élément de formalisme. Ceci est un exemple d'unification des formalisations, mais l'unification n'est pas toujours possible. Par exemple, la formalisation sous forme de réseau sémantique présente l'avantage de montrer d'un seul coup d'œil le nombre de sémantèmes qui la composent, ce qui ne serait pas

possible dans une définition DEC : c'est une limitation inhérente à cette dernière formalisation.

Une tentative d'uniformisation des formalisations aurait pour résultat :

- 1) une plus grande cohérence entre les formalisations ;
- 2) une facilitation dans la transition, automatique ou non, d'une formalisation à une autre ;
- 3) une amélioration du formalisme du réseau sémantique, qui se veut la formalisation idéale de la RSém.

Ainsi, dans le reste de ce chapitre, nous tenterons d'exposer les différences entre ces formalisations. Dans la section 2, nous verrons les disparités ayant trait à la structure sémantique (SSém), puis, dans la section 3, nous traiterons de celles qui concernent la structure communicative (SSém-Comm). Finalement, dans la section 4, nous présenterons un tableau récapitulatif contenant la synthèse de toutes les disparités discutées. Rappelons que c'est au chapitre 4 que nous proposerons des solutions pour uniformiser davantage les formalisations.

## **2 Disparités entre les trois formalisations ayant trait à la structure sémantique (SSém)**

La description des disparités nécessitera parfois l'approfondissement de certaines notions et l'éclaircissement de certaines questions, ce que nous tenterons de faire chemin faisant.

### ***2.1 Composante générique et nœud communicativement dominant***

Dans le réseau sémantique, on identifie le nœud communicativement dominant global (NDg). Or, parce qu'il est un nœud unique par définition, ce nœud ne correspond pas nécessairement à la composante générique dans son entièreté. Par exemple, si la composante générique est 'ensemble de personnes', le NDg du réseau sera seulement 'ensemble' et, si c'est 'outil de communication', le NDg sera seulement 'outil'. Dans la BDéf, la composante générique est explicitement identifiée et contient l'ensemble des sémantèmes le composant et non seulement le sémantème qui est le NDg. Par exemple, dans la définition BDéf de OURAGAN1, 'tempête d'extrême violence' est identifié comme étant la composante générique. Dans le réseau sémantique correspondant, seul le

sémantème ‘tempêteI’ est marqué comme ayant une importance particulière dans le sémantisme de la définition puisqu’il est le NDg de la définition. On voit qu’il n’y a pas identité entre le concept de NDg et celui de composante générique. Dans la définition DEC, ni l’un ni l’autre n’est identifié.

Polguère (1990), qui introduit la notion de nœud dominant (ND) dans la littérature, dit bien que la notion de nœud dominant global (NDg) pour une SSém d’énoncé a été inspirée par la notion de composante générique dans les SSém de lexie. Ainsi, si on accepte que la composante générique ne soit pas tout à fait équivalente au NDg, on peut se demander s’il est pertinent d’identifier le NDg dans la SSém d’une lexie en plus de la composante générique. Par contre, pour les nœuds dominants de sous-réseaux, il nous semble naturel que chacune des structures (sous-réseaux de la SSém et sous-réseaux de la SSém-Comm) gravite autour d’un sémantème plus important que les autres et qu’ils soient identifiés, que ce soit pour une SSém de phrases ou de lexies. Ainsi, pour qu’il y ait une cohérence entre les différentes structures, il faut conserver le NDg dans la RSém des lexies. Quant à la composante générique, puisqu’elle est au cœur de la notion de définition dans la LEC, il n’est pas question pour nous de ne pas l’identifier de façon explicite. Par conséquent, dans les formalisations proposées au chapitre 4, nous trouverons une indication explicite de la composante générique et du NDg dans chacune des formalisations.

## 2.2 Composantes périphériques

### 2.2.1 Séparation en blocs définitionnels élémentaires

Le terme de *bloc définitionnel élémentaire* appartient à la terminologie de la BDéf. Ce concept lexicologique, c’est-à-dire une subdivision de la structure sémantique en configuration de sémantèmes en fonction de leur rôle informationnel (le type de spécificité apporté à la composante générique, voir la section 2.3.2.1 du chapitre 2), devrait également être utilisé dans les autres formalisations. Nous utilisons le terme *bloc définitionnel élémentaire* ou *bloc élémentaire* pour parler de ces configurations de sémantèmes, et ce, pour toutes les formalisations traitées dans ce travail, même si, dans un réseau sémantique, par exemple, il ne s’agit pas à proprement parler d’un bloc, mais bien d’un sous-réseau d’un certain type.

L'identification de blocs définitionnels élémentaires constitue un des éléments qui distinguent nettement le format de la BDéf des autres formalisations présentées. Il n'y a pas d'équivalent manifeste aux blocs élémentaires dans la formalisation en réseau sémantique ni dans la définition lexicographique. Cette division en blocs permet d'explicitier l'organisation du sens d'une lexie. Il est de notre avis que leur absence dans la RSém des lexies est un manque à pallier. Ainsi, il faudra trouver comment ajouter les rôles informationnels de façon élégante et claire.

Les blocs élémentaires et leurs en-têtes, qui indiquent le rôle informationnel du bloc, ne sont pas que des métadonnées destinées à aider le lexicographe ou à faciliter la lecture d'un dictionnaire théorique. Ils font partie intégrante de l'organisation du sens et sont essentiels à notre capacité à utiliser les lexies. Plusieurs auteurs (Wierzbicka, 1985; Iordanskaja & Mel'čuk, 1990; Pustejovski, 1995; Padučeva, 1997; Mel'čuk & Wanner, 2001; Milićević, 2008) soulignent le besoin de structurer les définitions pour expliciter leur organisation interne.

Certains pourraient penser qu'il n'est pas nécessaire d'explicitier les blocs élémentaires à l'aide d'un nouvel élément de formalisme dans le réseau sémantique puisque le rôle informationnel pourrait être retrouvé automatiquement à partir de sémantèmes identifiant ce rôle. Par exemple, dans le réseau sémantique de PEIGNE1a (Réseau sémantique 2), le nœud 'but' donnerait une indication quant à la présence d'un bloc élémentaire ayant pour rôle informationnel /\*But\*/. Cependant, comme le notent Altman & Polguère (2003 : 13) en parlant de la BDéf :

Nous avons réalisé qu'il était important d'identifier clairement dans une (sic) en-tête le rôle informationnel de chaque composante plutôt que de se fier à la présence de tel ou tel sémantème (par exemple ici, '[dans le] but') dans le bloc en question. En effet, le rôle informationnel n'est pas toujours explicitement indiqué dans une définition par une composante spécifique, le même rôle pouvant être modélisé de diverses façons en fonction des sémantèmes spécifiques devant être introduits dans la composante en question. La nécessité d'avoir une bonne paraphrase prime, dans les définitions analytiques du DEC, sur celle d'encoder explicitement le rôle informationnel de chaque composante.

Bien sûr, on peut se questionner à savoir si cette remarque s'applique de la même façon aux réseaux sémantiques. Nous croyons qu'il est préférable d'identifier explicitement le rôle informationnel de chaque bloc élémentaire dans le réseau comme dans les autres formalisations pour éviter toute ambiguïté et pour expliciter la structure sémantique des lexies décrites. Le fait de repérer un sémantème qui identifie le rôle du bloc ne permettrait pas d'en déterminer la portée ni de savoir exactement quelle configuration de sémantèmes serait contenue dans le sous-réseau correspondant au bloc élémentaire. Nous verrons des exemples de réseaux dans lesquels les blocs élémentaires sont étiquetés dans nos propositions de formalisations au chapitre 4.

### 2.2.2 L'ordonnement des composantes sémantiques

Il nous faut maintenant expliquer en quoi l'ordre des composantes est important et pourquoi il faut le représenter dans nos formalisations. Après tout, la formalisation de base de la TST, le réseau sémantique, n'implique à la base aucun ordre de lecture. Nous présenterons trois arguments à cet effet.

Premièrement, pour Wierzbicka (1977 : 171), ce qui est en cause dans l'organisation d'une définition n'est pas la linéarité, elle-même, mais l'ordre temporel exprimé par cette linéarité. Pour Polguère (1992 : 135), « il n'y a pas d'accès possible à la perception intuitive d'un sens encodé dans une représentation quelle qu'elle soit sans le recours à un ordre temporel. » À partir de cet état de fait, il construit la règle suivante :

« Une représentation sémantique pourra prendre la forme d'une structure non linéaire pour peu que cette structure possède une hiérarchie interne impliquant un ordre temporel dans la lecture de la représentation. » (Polguère, *Ibid.*)

Deuxièmement, l'ordre de lecture des composantes est important pour expliciter le fait que certaines composantes sont subordonnées à d'autres composantes, en ce qui concerne la pertinence sémantique :

#### **Règle 6 : Règle d'ordonnement par pertinence sémantique**

« La composante exprimant une caractérisation ' $\sigma_1$ ' sémantiquement plus pertinente qu'une autre caractérisation ' $\sigma_2$ ' doit précéder la composante exprimant ' $\sigma_2$ '.

[...]

NB: Dans un réseau sémantique, il n'est évidemment pas possible d'utiliser la Règle 6, puisque les éléments d'un réseau ne sont pas soumis à l'ordre linéaire. Dans la représentation sémantique véritable, il faudra introduire une technique appropriée explicite pour marquer le degré de pertinence sémantique.» (Mel'čuk *et al.*, 1995 : 107).

À ce jour, cette problématique n'a pas été résolue. Il s'agit d'indiquer la pertinence sémantique dans le réseau pour suggérer un ordre de lecture. Cependant, dans la théorie Sens-Texte, l'ordre des symboles ne peut jamais être utilisé pour donner de l'information dans les représentations sémantiques ou syntaxiques ; tout ce qui doit être exprimé doit l'être à l'aide de symboles explicites (Mel'čuk, 1988b : 16).

Troisièmement, et cet argument rejoint un peu le précédent, il nous semble logique de placer des composantes optionnelles ou faibles à la suite des composantes qui ne le sont pas.

Pour nous, l'argument de la perception intuitive du sens est le plus fort, mais, bien sûr, l'importance de cet argument s'estompe lorsqu'on a affaire à un réseau destiné au traitement automatique du langage. En fait, il faut faire plus de descriptions explicitant la structure interne des représentations pour que l'ordre des composantes puisse être étudié en tant que tel et que des règles spécifiant son fonctionnement puissent être extraites.

Pour une phrase, l'opposition communicative thème/rhème, dans la SSém-Comm, donne certaines indications quant à l'importance relative de certaines composantes, ce qui constitue un des critères de leur ordonnancement. Pourtant, pour la RSém d'une lexie, l'importance communicative ne semble pas fournir suffisamment d'information pour reproduire dans un réseau l'équivalent de l'ordre de chacun des blocs définitionnels élémentaires dans une définition BDéf.

#### *2.2.2.1 Différents traitements dans les formalisations*

Une définition DEC, parce qu'elle utilise un langage très proche du langage naturel, impose que chaque élément de sens qui la compose soit placé dans un ordre déterminé, en

respectant ce qui est énoncé dans la Règle d'ordonnement par pertinence sémantique de l'ILEC, citée plus haut.

Aussi, dans une définition de type BDéf, un ordre de lecture est imposé puisque, du haut vers le bas de la définition, on lit chacun des blocs définitionnels élémentaires l'un à la suite de l'autre et chacune des composantes l'une à la suite de l'autre. Par contre, comme cela a été mentionné ci-dessus, la TST n'accepte pas que l'ordre linéaire soit porteur d'information dans une RSém. Dans ce cas, on pourrait penser que la numérotation de chacune des composantes dans une définition BDéf explicite cet ordre. Or, nous sommes d'avis que ces numéros de composantes ont seulement une fonction identificatrice qui permet de faire référence à une composante dans une autre composante (par ex. 2 : X croit que \*1) et n'ont pas d'autre valeur théorique. Ils marquent l'ordre, mais cet ordre n'est pas sous-tendu par une logique qui ordonnerait les composantes entre elles. Du moins, cette logique n'est explicitée nulle part.

Comme nous l'avons vu au chapitre 2, dans le réseau sémantique, l'ordre de lecture est absent, sauf pour ce qui est du NDg. L'absence d'un ordonnancement des composantes rend la lecture et l'interprétation de la SSém beaucoup moins intuitive et donc plus difficile.

### **2.2.3 Organisation en blocs et ordre de lecture des blocs dans les formalisations à l'étude**

Nous croyons que la pertinence sémantique, qui dicte l'ordre de lecture des composantes, est déterminée par un schéma de définition associé au champ sémantique de la lexie.

#### **2.2.3.1 Réseaux sémantiques**

Dans la BDéf, les différentes composantes sémantiques présentes dans une définition sont clairement identifiées. Elles sont isolées en blocs définitionnels élémentaires et étiquetées d'un rôle informationnel. Il est possible d'insérer les mêmes divisions dans les réseaux sémantiques. Par contre, l'ordre de lecture des composantes, inévitablement présent dans une formalisation linéaire comme la définition lexicographique ou la définition BDéf, n'a pas son équivalent formel dans le réseau sémantique.

#### **2.2.3.2 La définition DEC**

La définition DEC n'est pas sémantiquement structurée. Les sémantèmes semblent tous revêtir la même importance dans la définition. Ils ne sont pas identifiés comme appartenant

soit à la composante générique, soit aux composantes périphériques de la lexie décrite. Aussi, les composantes périphériques ne sont pas séparées entre elles en fonction de leur rôle informationnel, c'est-à-dire en fonction du type de sémantisme qu'elles apportent à la composante générique.

Ainsi, nous considérons que la définition lexicographique forme un tout presque homogène au niveau de la SSém. À ce tout se superpose l'opposition communicative posé/présumé, qui appartient à la SSém-Comm.

### 2.3 *Explicitation des liens de polysémie*

#### 2.3.1.1 *Différents traitements dans les formalisations*

Comme nous l'avons vu plus haut, dans le DEC, on indique de façon explicite les ponts sémantiques de nature métaphorique. La composante sémantique qui contient le pont sémantique est mise entre crochets « [...] » et débute par le sémantème 'comme si'. Par contre, les ponts sémantiques de nature métonymiques, eux, sont introduits par différents sémantèmes, comme le montrent ces exemples tirés de l'ILEC (Mel'čuk *et al.*, 1995 : 164) :

- 'localisé à...' (*La France(L2) y était opposée – la population de la France(L1)*<sup>6</sup>)
- 'provenant de...' (*beaujolais(L2) - vin du Beaujolais(L1)*)
- 'destiné à...' (*jambe(L2) de pantalon – jambe(L1)*)
- 'contenu dans...' (*tasse(L2) – tasse(L1) de farine*)
- 'fabriqué de...' (*un verre(L2) - fabriqué de verre(L1)*)
- '... qui a...' (*un cœur généreux(L2) – qui a un cœur généreux(L1)*)

Dans tous les cas, la lexie de base (L1) est incluse dans la définition de la lexie « métonymique » (L2). Ainsi, dans les paires de définitions suivantes, on sait qu'il y a un lien de polysémie grâce à la présence du sémantème en gras, mais c'est grâce à la composante soulignée que l'on sait qu'il s'agit d'une relation métonymique et que l'on connaît le type de métonymie dont il s'agit.

---

<sup>6</sup> Le soulignement double indique ce qui a été ajouté à l'extrait original de l'ILEC.

L1 : VERRE-**I.1** : substance fabriquée, dure, cassante et transparente, de structure vitreuse

L2 : VERRE-**II.1** : récipient à boire typiquement fabriqué de verre-I.1

L1 : TASSE-**1a**: petit récipient à anse ou à oreilles servant à boire (généralement des boissons chaudes)

L2 : TASSE-**1b**: contenu d'une tasse-**1a**

Dans ces circonstances, la métonymie est difficile à repérer dans la définition et un type de relation sémantique entre les lexies (contenu-contenant, proximité, etc.) en particulier l'est encore davantage, puisque les formules introduisant la métonymie peuvent subir des variations stylistiques d'une définition à une autre ('localisé à', 'se trouvant à', etc.).

Ceci est d'autant plus vrai si l'on accepte la typologie de Barque (2008), qui inclut d'autres types de lien de polysémie que la métaphore et la métonymie. Dans l'état actuel des choses, on sait seulement que, si on ne trouve pas le sémantème 'comme si', alors on n'est pas en présence d'une métaphore. Par conséquent, il est de notre avis que, tout comme la métaphore, les liens métonymiques devraient être identifiés de façon explicite dans les définitions du DEC. Malheureusement, le lien de métonymie ne se résume pas bien en une expression-clé comme le « comme si » de la métaphore.

En ce qui concerne le réseau sémantique, en tant que formalisation du sens lexical, aucun formalisme n'a été développé pour expliciter les liens de polysémie entre les lexies d'un même vocable, sauf pour ce qui est de la présence du sémantème 'comme si', qui indique qu'on est en présence d'une métaphore. Dans la BDéf, les deux types de ponts sémantiques – métaphoriques et métonymiques – sont indiqués dans l'en-tête du bloc. Nous proposerons de nouveaux formalismes au chapitre 4, et ce, pour chacune des trois formalisations.

## 2.4 *Contraintes sur les actants*

### 2.4.1 *Description de la notion*

Chaque actant, instancié par une variable dans les formalisations, peut être contraint par un sémantème ou une configuration de sémantèmes qui précise quels lexèmes pourront remplir la position actancielle identifiée par la variable en question. Cette contrainte s'appelle

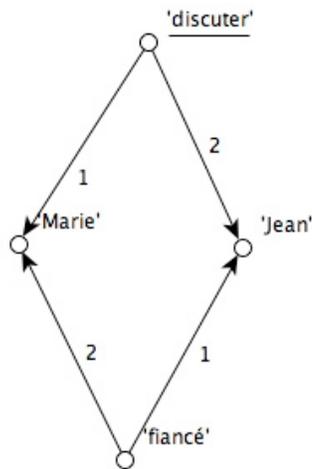
*contrainte actancielle*. Cette dernière n'ajoute pas de sens à une définition. Elle ne fait que restreindre les valeurs que peut prendre une variable actancielle. Elle ne se trouve que dans la définition de la lexie ; elle n'est plus pertinente lorsque la lexie est utilisée dans une phrase, et elle n'apparaît donc pas dans la SSém de cet énoncé. À titre d'exemple, reportons-nous à la définition suivante :

FIANCÉ

'X est le ~ de Y' = 'homme X tel qu'il est formellement engagé à épouser la femme Y'

(17) *Marie discute avec Jean, son fiancé.*

(18)



Les sémantèmes 'Marie' et 'Jean' remplissent les positions actanciennes de la lexie FIANCÉ. Les contraintes actanciennes 'homme' et 'femme' ne sont plus pertinentes pour la SSém de l'énoncé, sauf pour vérifier la bonne formation de la SSém.

Il est à noter que les contraintes sur les actants agissent toujours comme un présupposé. Soit la phrase :

(19) *Il n'a pas bu l'ettouhnav.*

Malgré que le verbe BOIRE<sub>I.1</sub> soit nié, on affirme toujours que *il* réfère à un individu ou un animal et que *ettouhnav* est un liquide. C'est parce que, dans la définition de BOIRE<sub>I.1</sub> [X boit Y], 'X' porte une contrainte actancielle 'individu ou animal' et 'Y' porte une contrainte 'liquide'.

#### 2.4.2 Traitement de l'identification des contraintes sur les actants dans les différentes formalisations

Dans les réseaux sémantiques, les contraintes sur les actants, lorsqu'elles sont présentes, ne sont pas indiquées par un formalisme particulier. Elles sont plutôt introduites par le sémantème 'être' (par ex., 'X' <-1- 'être' -2-> 'animal'). Dans le DEC, les contraintes sur les actants sont parfois mises dans une structure de subordonnée relative :

VEUVE (DECFC-IV : 101)

X est veuve = 'X, qui est une femme [...], a perdu son mari'

Dans la BDéf, un bloc à part du corps de la définition leur est dédié, le bloc des contraintes sur les actants.

Une variable actancielle n'est pas obligatoirement accompagnée d'une contrainte actancielle. Dans les faits, l'absence d'une contrainte actancielle explicite devrait être interprétée comme étant l'indication d'une valeur par défaut, très générale comme 'quelque chose', correspondant à l'étiquette sémantique *quelque chose*, qui domine la hiérarchie d'étiquettes.

#### 2.5 Les connecteurs

Dans une représentation en langage naturel comme la définition DEC, on est obligé d'utiliser des éléments syntaxiques sémantiquement vides. C'est, par exemple, le cas du 'et' dans la définition de OURAGAN $\mathbf{I}$  :

OURAGAN $\mathbf{I}$

'Tempête $\mathbf{I}$  d'extrême violence [...] qui tend à avoir des effets destructeurs sur Z se trouvant dans Y et qui se déplace vers W.'

Ces éléments sémantiquement vides ne se trouvent pas dans le réseau sémantique, ni dans la BDéf. Nous jugeons que les connecteurs, lorsqu'ils sont sémantiquement vides, devraient être marqués pour les distinguer de leur contrepartie sémantiquement pleine.

#### 2.6 Composantes faibles et optionnelles

Les composantes faibles étant rares, nous ne connaissons pas bien leur traitement dans les formalisations. Nous avons accès à quelques exemples dans le format de la définition DEC

puisque'il en est question dans l'ILEC. Ainsi, nous nous concentrerons sur la composante optionnelle, mieux documentée, en gardant en tête que la composante faible sera probablement traitée de façon très similaire.

Dans la BDéf, les composantes optionnelles sont mises entre parenthèses, tout comme dans les définitions DEC. Par exemple, dans la définition BDéf de CHICHEI.1 [*— Chiche que tu ne sauteras pas!*] :

```
[...]
/*But*/
3: *1 dans_le_but *4
4: locuteur montrer à *5 que *6
5: et
5.1: interlocuteur
5.2: (autre {individu})
6: ou
6.1: interlocuteur incapable de Z
6.2: interlocuteur avoir peur de Z
[...]
```

**Figure 15 : Extrait de la définition BDéf de CHICHEI.1**

Ce bloc définitionnel élémentaire se lit comme suit : ‘dans le but de montrer à son interlocuteur (et à d’autres individus) que l’interlocuteur est incapable ou a peur de Z-er’. La composante ‘autres individus’ est optionnelle puisque le locuteur pourrait utiliser la lexie CHICHEI.1 en l’absence d’autres individus que son interlocuteur.

Pour ce qui est du traitement des composantes optionnelles dans le réseau sémantique, il est mentionné dans la littérature de la LEC que le sous-réseau correspondant à une composante optionnelle est encerclé dans le réseau. Ce formalisme nous semble insuffisant puisque l’encerclément est utilisé pour identifier des sous-réseaux de tous types.

### **3 Les disparités qui concernent la structure communicative (SSém-Comm)**

#### **3.1 La thémativité**

Comme nous l’avons expliqué précédemment (chapitre 2, section 5.1), la thémativité n’est pas traitée dans ce travail.

### 3.2 *La locutionnalité*

Dans chacune des formalisations, les sémantèmes marquant la valeur illocutoire de la lexie devraient être identifiés comme tels. Il s'agit principalement des sémantèmes indiquant les signalitifs puisque les lexies « communiquées » représentent le cas non marqué et que les formalismes pour indiquer les performatifs ne sont pas développés.

Pour le réseau sémantique, il faut identifier les sémantèmes 'locuteur' et 'signaler'. Pour le DEC, la composante 'le locuteur signale que' introduit les signalitifs. Finalement, pour la BDéf, il s'agit de la même composante en « langage BDéf » : 'locuteur signaler...'.

### 3.3 *La présupposition*

Les éléments de formalisme exprimant la présupposition dans les trois formalisations étudiées sont différents. Dans le réseau, on utilise un sous-réseau ombragé. Dans la définition DEC, on marque la présupposition avec « |[...]| ». Dans la BDéf, c'est dans l'en-tête du bloc qu'on marque la composante présupposée.

## 4 **Tableau synthèse des disparités entre les formalisations**

Les concepts lexicologiques qui ont été identifiés comme étant présents dans les trois formalisations en utilisant des formalismes semblables ne sont pas répétés dans ce tableau. Chacune des disparités présentées dans le tableau a été expliquée dans ce chapitre. Les cases grisées identifient les éléments de formalisme inexistant ou insuffisant pour représenter un concept lexicologique donné.

	<b>Réseau sémantique</b>	<b>Définition DEC</b>	<b>Définition BDéf</b>
Éléments de formalisme SSém			
Identification de la composante générique	NON	NON	OUI bloc central
Identification du NDg	OUI <u>soulignement</u>	NON	NON
Présence des blocs définitionnels	NON	NON	OUI
Typages des blocs définitionnels	NON	NON	OUI /*En-tête du bloc*/

	<b>Réseau sémantique</b>	<b>Définition DEC</b>	<b>Définition BDéf</b>
Ordre des composantes	NON	OUI ordre linéaire langagier	OUI et NON <sup>7</sup> Numérotation : *1, *2, etc.
Explicitation des liens métaphoriques	OUI 'comme si'	OUI [comme si ...]	OUI /*[lien_métaphorique]*/
Explicitation des liens métonymiques	NON	NON	OUI /*[lien_métonymique]*/
Explicitation des relations sémantiques entre les actants	OUI flèches	NON	OUI Bloc des relations sémantiques
Composantes optionnelles	OUI mais insuffisant sous-réseau encerclé	OUI (composante opt.)	OUI (composante opt.)
Composantes faibles	OUI mais insuffisant sous-réseau encerclé	OUI (composante faible) <sub>condition</sub>	NON <sup>8</sup>
Explicitation du typage d'actant	NON	NON	OUI Bloc de typage des actants
Distinction/absence des connecteurs sémantiquement vides	OUI Ils sont absents.	NON	OUI Ils sont absents.
Éléments de formalisme SSém-Comm			
Présupposés	OUI sous-réseau ombragé	OUI avant la double barre    ou   ...	OUI /*[présumé]*/
Explicitation des sémantèmes exprimant la locutionnalité	NON	NON	NON

**Tableau I : Tableau résumant les éléments de formalisme associés à chacun des concepts lexicologiques analysés, et ce, pour chacune des trois formalisations traitées dans ce travail**

<sup>7</sup> Rappelons que nous avons émis des réserves à ce sujet puisque la numérotation dans la définition BDéf ne repose sur aucune logique explicitée.

<sup>8</sup> Nous ne disposons d'aucun exemple et il n'en est pas question dans la littérature de la BDéf.

## CHAPITRE 4 Proposition d'améliorations à apporter aux trois formalisations de la RSém

### 1 Introduction

Maintenant que nous avons identifié les disparités entre les formalisations de la représentation sémantique dans l'approche Sens-Texte, nous sommes en mesure de proposer des améliorations ou des normalisations aux formalismes utilisés afin de les compléter et de les uniformiser.

Nous nous donnons six lignes directrices à respecter :

- 1) représenter tous les éléments requis pour une utilisation adéquate de la lexie L, en accord avec le principe d'exhaustivité :

« Dans un DEC, l'article d'une lexie L inclut tous les renseignements lexicographiques nécessaires, d'une part, pour bien utiliser L elle-même, et d'autre part, pour pouvoir trouver les autres lexies L<sub>1</sub>, L<sub>2</sub>, etc. sémantiquement liées à L. » (Mel'čuk *et al.*, 1995 : 42)

En ce qui nous concerne, nous ne considérons que la zone sémantique. Nous ne tenons pas compte ici des informations phonologiques, des connotations, de la combinatoire syntaxique et lexicale.

- 2) représenter les mêmes concepts lexicographiques d'une formalisation à l'autre, lorsque cela est possible.
- 3) utiliser les mêmes éléments de formalisme d'une formalisation à l'autre, lorsque cela est possible.
- 4) marquer les concepts lexicologiques appartenant à des structures distinctes (SSém et SSém-Comm) de façons différentes.
- 5) conserver un certain degré de lisibilité.
- 6) éviter la redondance dans l'information lexicographique donnée, c'est-à-dire que, autant que possible, chaque concept lexicographique ne devrait être représenté que par un seul élément de formalisme.

La seule dérogation permise a trait au point 1 et concerne les définitions destinées aux apprenants. En effet, il ne serait pas avantageux, dans ce cadre particulier, que les définitions soient surchargées de formalismes lourds à apprendre. Cela dit, nous considérons qu'aucune des formalisations étudiées n'est destinée aux apprenants.

À la section 2, nous présenterons un outil de description très puissant: le schéma de définition. Nous verrons les propositions de modifications qui concernent le réseau sémantique, la définition DEC et la définition BDéf, aux sections 3, 4 et 5 respectivement. À la section 6, nous ferons quelques remarques sur les formalismes de la SSém-Comm. Finalement, nous soumettrons une nouvelle version du tableau comportant les disparités entre les formalisations, présenté au chapitre 3, afin d'y ajouter les éléments de formalismes proposés.

## **2 Les schémas de définitions**

Le schéma de définition est un outil fort intéressant pour le travail lexicographique, qui n'a pas été formalisé de façon stricte dans la TST. Plusieurs auteurs mentionnent l'utilisation de schémas de définitions (Rey, 1977; Wierzbicka, 1985; Pustejovski, 1995; Iordanskaja & Mel'čuk, 1990; Padučeva, 1997; Milićević, 2008; Barque, 2008). Il est aussi appelé *patron de définition*, *schéma généralisé* ou *format standard* par certains auteurs.

Un schéma de définition est une liste structurée des types de composantes que devrait contenir la définition d'une lexie appartenant à un champ sémantique donné. Chacune des composantes est désignée par son rôle informationnel. Le schéma n'est pas rigide ; les composantes qu'il suggère peuvent ne pas se trouver dans la définition et des types de composantes non prévues par le schéma peuvent aussi s'y trouver. Le schéma de définition est ordonné et hiérarchisé. Selon certains auteurs, il peut aussi inclure des éléments de la SSém-Comm, telle la présupposition.

Dans la littérature, les schémas sont vus comme une aide à la rédaction de dictionnaire, à la consultation de dictionnaire ou à une réflexion approfondie sur l'organisation interne du sens. Toutefois, aucun lexicographe ne propose de compilation de l'ensemble des schémas qui ont été utilisés pour les différentes définitions qu'ils ont construites. Aussi, les

dictionnaires « théoriques » comme le DEC ne présentent pas les schémas associés à chacun des champs sémantiques qui y sont traités. Aussi, Iordanskaja & Mel'čuk (1990) expriment des réserves quant à la pertinence d'identifier les composantes d'une définition à l'intérieur de la définition elle-même et croient qu'il vaudrait mieux avoir seulement un schéma de définition au niveau du champ sémantique. Cela nous semble être un peu le problème de la poule et de l'œuf puisqu'on a besoin d'identifier les différents types de composantes sémantiques dans les définitions afin de pouvoir construire un schéma de définition pour le champ sémantique.

Dans le cas de la BDéf, on expose la structure du sens de chacune des lexies avec des blocs définitionnels, munis d'un en-tête indiquant leur rôle informationnel. On espère pouvoir en extraire des schémas de définitions une fois qu'un certain nombre de lexies d'un même champ seront décrites. Nous présenterons maintenant quelques schémas de définition que l'on retrouve dans la littérature.

### 2.1 Schéma proposé par Iordanskaja et Mel'čuk

Iordanskaja & Mel'čuk (1990 : 17) proposent un schéma de définition pour le champ sémantique des ÉMOTIONS:

**Présumé:** ...

**Posé:**

**Évaluation** (de la situation, par X): ...

**Émotion** (de X):

**Caractérisation** (en fonction des dimensions pertinentes): ...

**Cause** (= les conditions) : ...

**Conditions prototypiques** : ...

**Effet** (de l'émotion de X sur X): ...

Schéma de définition 1 : Schéma de définition de Iordanskaja & Mel'čuk (1990 : 17)

Voici la définition de БОЈАТ'СЈАИ.1а ('avoir peur') correspondant au schéma ci-dessus. Les en-têtes de blocs se trouvent dans la colonne de gauche. La colonne de droite constitue la définition comme telle.

BOJAT'SJAL.1a	X boitsja Y-a 'X a peur de Y'=
<b>Posé</b>	
<b>Évaluation</b>	'X s'attend à ce que Y cause quelque chose de non désirable pour X
	ou
<b>Émotion</b>	X est ou a la propriété d'être dans
<b>Caractérisation</b>	un état émotionnel déplaisant en lien avec Y
<b>Cause</b>	qui est causé par lesdites attentes de X,
<b>Conditions prototypiques</b>	cet état est comme celui causé par l'attente de quelque chose de dangereux,
<b>Effet</b>	cette attente et/ou état a pour effet que X tende à éviter Y'.

**Définition 16 : Définition de BOJAT'SJAL.1a**

Iordanskaja & Mel'čuk (1990 : 17) spécifient qu'il s'agit d'un schéma de définition maximal, certains champs sont facultatifs. Par exemple, la définition ci-dessus ne comprend pas de bloc « Présupposé ».

## 2.2 Schéma proposé par Padučeva

Padučeva (1997 : 62) parle de la définition lexicographique d'un énoncé. Pour la phrase *John killed a fly*, elle propose la définition suivante :

- I. Causateur: Jean a agi dans un certain But.
- II. Causation (contrôlée): (I) résulte en (III).
- III. Le Résultat coïncide avec le But: la mouche cesse de vivre.

Il s'agit bien d'une explicitation de la structure sémantique d'un énoncé, mais un énoncé n'a pas une structure préétablie due à son appartenance à une catégorie générale comme une lexie en a une parce qu'elle appartient à un champ sémantique. Toutefois, parce qu'on identifie et qu'on marque d'un en-tête les différentes composantes sémantiques mises en jeu dans la phrase, on peut extraire le schéma du verbe KILL (TUER) de cette « définition de phrase » et bâtir un schéma pour les LEXIES CAUSATIVES, dont le schéma serait quelque chose comme :

- I. Causateur: X a agi dans un certain But.
- II. Causation (contrôlée): (I) résulte en (III).
- III. Le Résultat coïncide avec le But.

Schéma de définition 2 : Schéma de définition de Padučeva (1997 : 62)

### 2.3 Schéma proposé par Mel'čuk et Wanner

Mel'čuk & Wanner (2001 : 15) proposent un schéma pour les noms de TEXTES QUI SONT LE RÉSULTAT D'UN ACTE DE LANGAGE PUBLIC (*names of texts resulting from public speech acts*) par exemple : APPEL, DÉCLARATION, PROCLAMATION, ANNONCE.

- 1. Composante générique
- 2. Posé
  - 2.1 Action centrale
  - 2.2 Acte de langage
  - 2.3 But (de l'acte de langage)
  - 2.4 Caractéristiques (de l'acte de langage)
- 3. Présuppositions
- 4. Actants

Schéma de définition 3 : Schéma de définition de Mel'čuk & Wanner (2001 : 15)

On parle de blocs de nature logique, communs à toutes les définitions (« Composante générique », « Posé », « Présuppositions », « Actants »). Le bloc « Posé » contient des sous-blocs de nature sémantique qui varient d'une définition à l'autre. Dans ce schéma de définition, il s'agit des points 2.1 à 2.4.

Voici un exemple de définition de la lexie allemande APPEL (en français, APPEL) qui est construite selon le schéma proposé ci-dessus :

APPEL [X, Y, Z]	
<b>1. Composante générique</b>	texte qui se veut accessible publiquement, par lequel
<b>2. Posé</b>	
<b>1.1 Action centrale</b>	X essaie de causer que Y fasse Z

<b>1.2 Acte de langage</b>	en communiquant à Y que X veut que Y fasse Z et en demandant à Y de faire Z
<b>1.3 But (de l'acte de langage)</b>	Pour que les émotions de X soient transmises à Y et pour causer que Y fasse Z
<b>1.4 Caractéristiques (de l'acte de langage)</b>	1. Z est suffisamment spécifié par rapport à l'état des choses $\alpha$ 2. le fait de communiquer et de demander a une teneur émotionnelle
<b>3. Présuppositions</b>	1. X croit que Y, où Y := {1.1. il y a un état des choses $\alpha$ qui concerne X, Y ou quelqu'un d'autre de façon vitale 1.2. $\alpha$ a ou aura des effets négatifs sur X, Y ou quelqu'un d'autre 1.3. pour éviter les effets négatifs de $\alpha$ l'action Z est nécessaire 1.4. Y peut et devrait faire Z} 2. X ressent des émotions vives à propos de Y
<b>4. Actants</b>	X = Communicateur : individu OU organisme officiel Y = Destinataire : individu important socialement OU collectivité importante socialement Z = But : action ou activité

En plus de la structure en blocs, deux dimensions s'ajoutent à l'organisation des définitions : les composantes fortes ~ faibles, présentées plus haut (chapitre 2, section 2.2.2.1), et les blocs primaires ~ secondaires, les blocs primaires pouvant être ignorés pour certaines applications (ce schéma de définition a été proposé pour une application de traduction automatique), leur contenu sémantique étant moins fort, moins essentiel, que celui des composantes primaires.

#### **2.4 Schéma proposé par Milićević**

Milićević (2008 : 4) présente ce qu'elle appelle la *structure de définition* pour les VERBES ÉVALUATIFS DE TYPE « DÉSA approuver ». Les en-têtes de bloc se trouvent dans la

colonne de gauche, mais les rôles informationnels apparaissent entre crochets au côté des composantes sémantiques, dans la colonne de droite.

DÉFINI = Forme propositionnelle	X = personne; Y = personne; Z = fait
DÉFINISSANT	
Présumé(s)	'X CROIT que Z a eu lieu/a lieu/peut avoir lieu' 'X CROIT que Y est responsable pour Z' ( 'X a l'influence/l'autorité sur Y' ) ( 'X est Z sont relativement proches' ) ( 'Y veut savoir ce que X pense à propos de Z' ) ( 'Z se trouve dans la sphère personnelle de X' )
Assertion	
1. Composante centrale	'X PENSE [attitude mentale] que Y/Z est BON/MAUVAIS [évaluation]'
2. Composante faible	( 'et X DIT [communication de l'évaluation] que Y/Z est BON/MAUVAIS' )
3. Composantes optionnelles	'X dit cela DANS LE BUT DE...[but de la communication]' 'X dit cela PARCE QUE...[motif de la communication]'

Schéma de définition 4 : Schéma de définition de Milićević (2008 : 4)

Il s'agit d'un schéma de définition. Nous trouvons très intéressant de lister ainsi les différentes composantes pouvant remplir chacune des « cases » du schéma de définition.

Voici la définition de CONDAMNER<sup>1I.1</sup> qui correspond à ce schéma. Nous avons déplacé les rôles informationnels dans la colonne de gauche.

CONDAMNER <sup>1I.1</sup>	
LA personne X ~ LA personne Y POUR LE fait Z	
<b>Présumé</b>	
	X CROIT que
	Z a eu lieu ou a lieu,
	Y est responsable pour Z.
<b>Assertion</b>	
<b>communication de l'attitude mentale</b>	X DIT que X PENSE
<b>évaluation</b>	que Z est très MAUVAIS
	parce que Z est contraire à certaines normes sociales ou morales
	et que Y est très MAUVAIS
	parce qu'il est responsable pour Z.
<b>but de la communication</b>	X dit cela DANS LE BUT DE
	nuire à Y
	ou de causer que Y change son comportement vis-à-vis de Z.

**Définition 17 : Définition de CONDAMNER<sup>1I.1</sup>**

## 2.5 Wierzbicka et le MSN

Wierzbicka utilise des en-têtes et des blocs dans certaines de ses définitions. Lorsqu'on lit une de ses définitions (voir celle de MUG plus haut à la Définition 5), on comprend la nécessité des en-têtes de bloc, qui rendent explicite l'organisation des sens contenus dans la définition. Wierzbicka avertit pourtant ses lecteurs que ces balises ne servent qu'à faciliter la lecture et qu'elles ne font pas partie de la définition en tant que telle (Wierzbicka, 1985 : 32). Par contre, elle affirme par la suite que l'utilisation de schémas de définitions est essentielle à l'interprétation de concepts, surtout lorsque ceux-ci sont très complexes. Ainsi, nous sommes d'avis que les en-têtes de blocs utilisés par Wierzbicka sont le reflet d'un schéma de définition que Wierzbicka juge en fait indispensable à l'interprétation du sens. Conséquemment, après avoir offert une définition de CUP et MUG, elle propose le schéma suivant pour les noms d'ARTEFACTS :

catégorie (une sorte de chose faite par des gens)  
 affectation [= angl. *purpose*] (relativement à la situation)  
 matériau  
 forme  
 taille

**Schéma de définition 5 : Schéma de définition de Wierzbicka (1985 : 52)**

Le fait qu'une définition débute par une catégorisation de l'objet à définir va presque de soi pour la plupart des lexicographes. Suit l'affectation, qui découle directement de la catégorie (une sorte de chose faite par des gens) puisqu'un objet est fabriqué par des gens dans un certain but. Ensuite, le matériau, la forme et la taille dépendent de l'affectation ayant motivé la fabrication de l'objet. Chacun des éléments de ce schéma équivaut à un bloc définitionnel élémentaire. Wierzbicka (1985 : 52-53) dit bien que ces schémas ne sont pas qu'un outil lexicographique :

Le schéma proposé ici ne se veut pas un simple outil technique pour construire des explications lisibles. Il s'agit d'une hypothèse sur la structure cognitive. Je présume que des concepts « simples » comme *cup* sont en fait formidablement complexes, comme l'indiquent mes explications excessivement longues, et ils doivent avoir des outils intégrés dans leur structure qui permettraient aux gens de les utiliser sans difficulté. Un schéma constant gouvernant l'ordre des composantes constituerait un parfait outil de ce genre. [Notre traduction.]

Elle insiste sur le fait que davantage de recherches sont nécessaires pour définir la structure que doivent suivre de tels schémas.

### **3 Modifications à apporter au réseau sémantique**

Il est important de noter que notre description des réseaux sémantiques est faite à partir de la forme « de base » du réseau, dépourvue de sous-réseaux sémantiques et communicatifs (au niveau de la SSém et au niveau de la SSém-Comm). Néanmoins, plusieurs auteurs ont proposé des formats plus structurés du réseau sémantique, nommons Dostie *et al.* (1999) et Polguère (2002a). Ces formats présentent plusieurs aspects dont nous nous inspirerons pour recommander des modifications aux formalisations.

### 3.1 Formalisations alternatives de réseau sémantique proposées par d'autres auteurs

Ainsi, Dostie *et al.* (1999) présentent une formalisation intéressante du sens de la lexie REPROCHER1, qui permet d'exposer la structure interne de la définition.

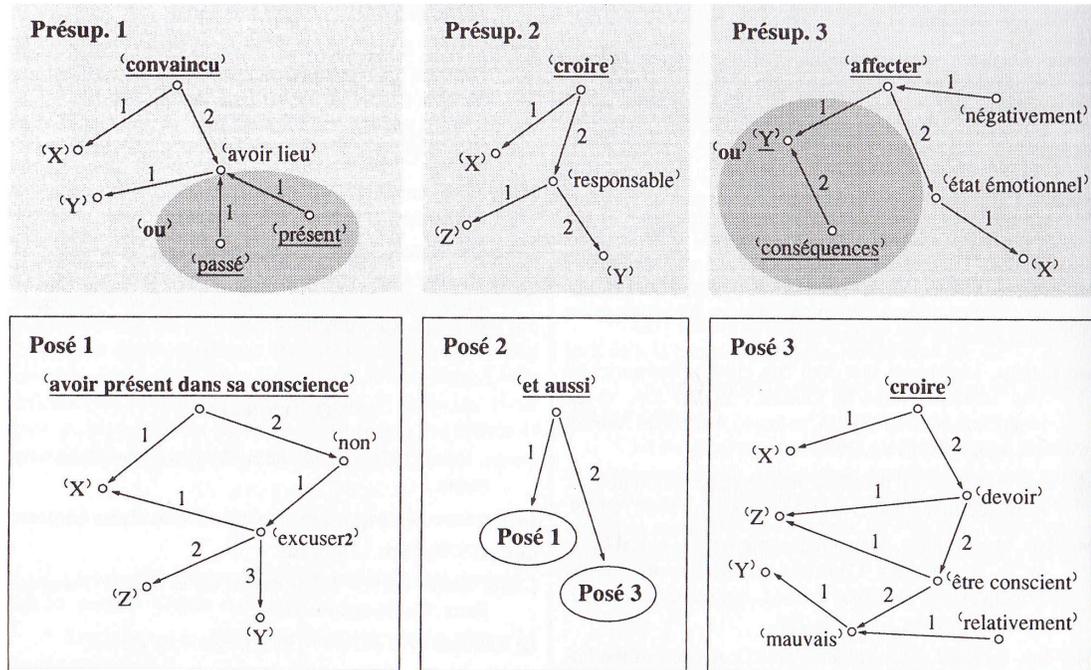


Figure 16 : Définition de REPROCHER1 : proposition de formalisation de Dostie *et al.* (1999)

Cette formalisation ou cette version du réseau sémantique présente l'avantage d'offrir un ordre de lecture et de subdiviser les sémantèmes en blocs en fonction de leur rôle informationnel. Par contre, il ne s'agit pas d'un graphe connecté, c'est-à-dire que les blocs ne sont pas connectés entre eux et qu'on ne voit pas d'interaction entre les blocs, sauf pour ce qui est des deux composantes liées par la conjonction 'et aussi' dans le bloc « Posé 2 ».

Aussi, la définition n'est pas structurée en fonction de la composante générique et des composantes périphériques. Le rôle informationnel des blocs n'est pas identifié, mais il pourrait facilement être ajouté. Finalement, ce format divise les blocs en fonction de l'opposition communicative posé/présumé et ne prévoit pas de façon d'identifier les autres oppositions communicatives, soit la thématique et la locutionnalité.

Polguère (2002a : 15) propose pour sa part cette formalisation :

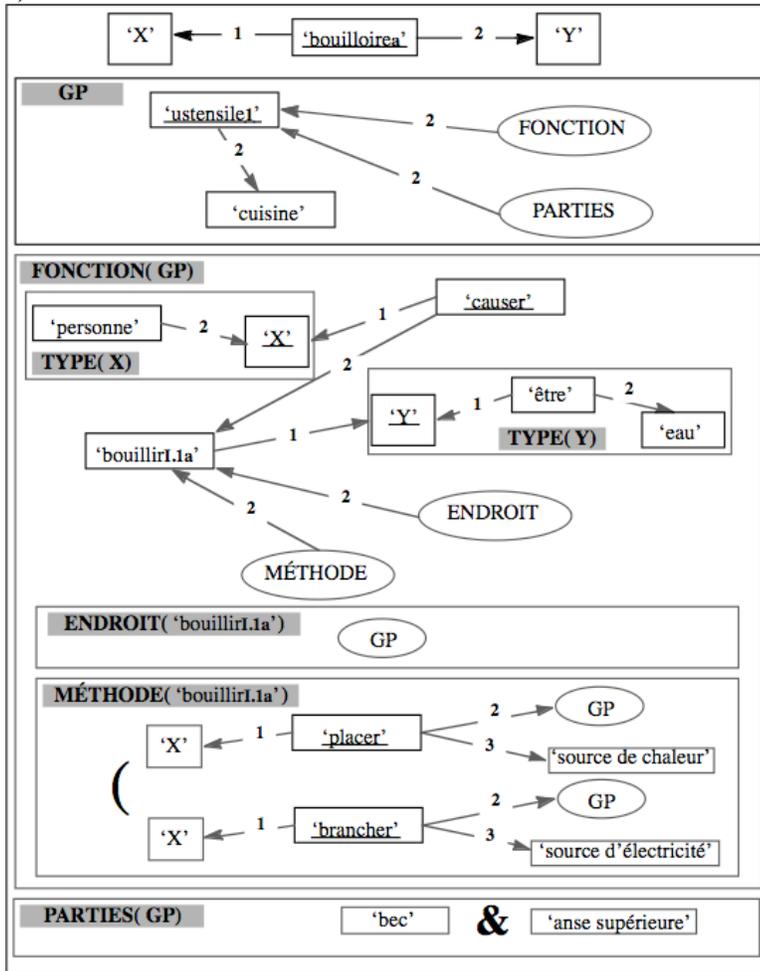


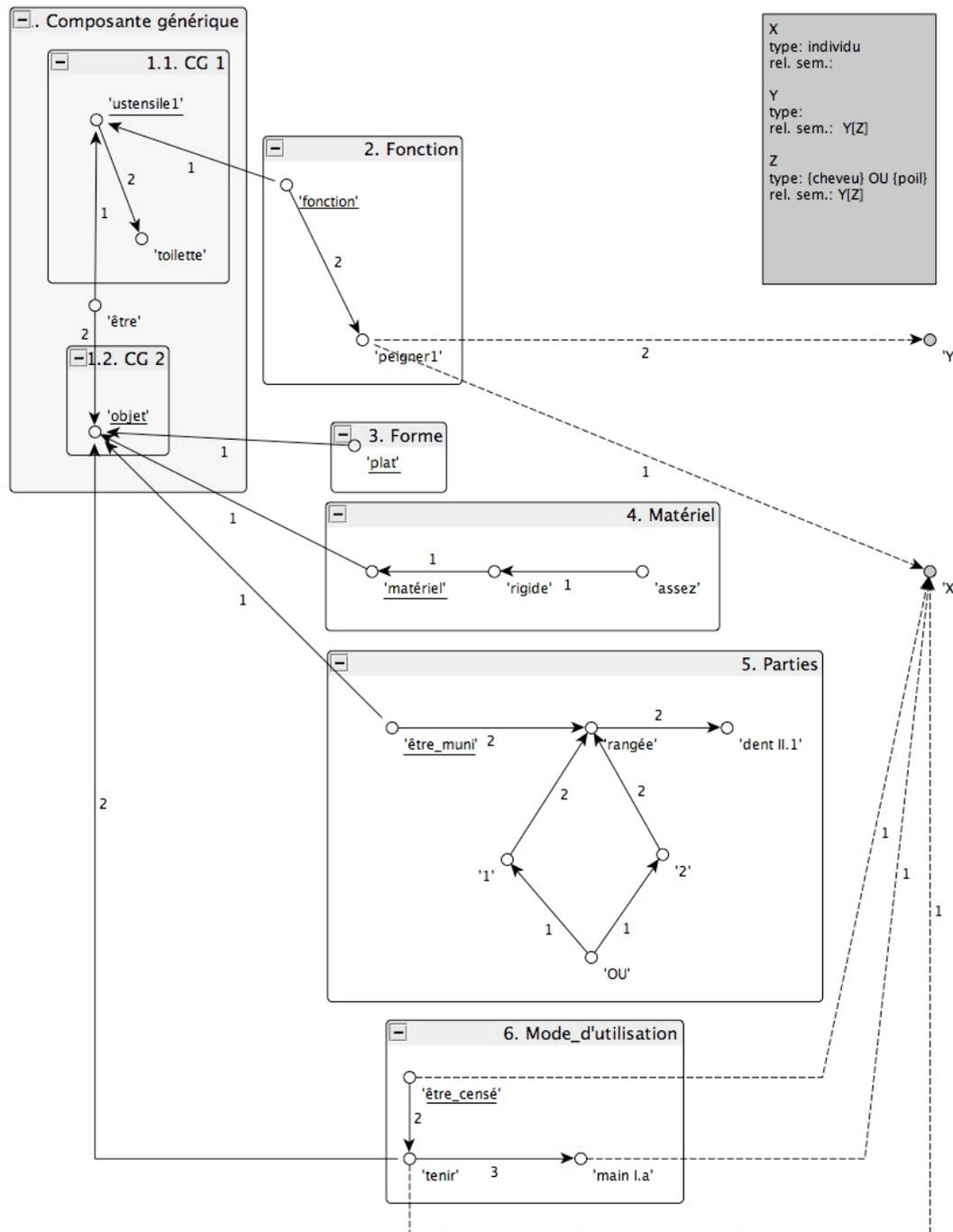
Figure 17 : Définition de BOUILLOIREA : proposition de formalisation de Polguère (2002a)

Cette version du réseau sémantique offre un ordre de lecture et une subdivision des sémantèmes en blocs en fonction de leur rôle informationnel et ceux-ci sont explicites. Pourtant, ce format n'est pas très facile à lire. Parce que les blocs renferment des références à d'autres blocs, on a tendance à les lire de façon récursive. Ainsi, le bloc ENDROIT contient le GP, qui lui contient le bloc FONCTION, qui contient, à son tour, le bloc ENDROIT. Par contre, dans cette formalisation, les liens entre les blocs sont très explicites.

### 3.2 Notre proposition de modifications à apporter au réseau sémantique

Nous proposons également de structurer le réseau sémantique afin d'explicitier sa structure interne. Pour ce faire, il est nécessaire de faire appel aux blocs définitionnels. Ces blocs seront étiquetés de leur rôle informationnel, à l'aide d'un en-tête. La structure du réseau doit également exposer la pertinence sémantique de chacune de ses composantes. Nous

présenterons un réseau sémantique remodelé à notre façon (Réseau sémantique 4), puis nous détaillerons ses particularités.



**Réseau sémantique 4 : Proposition d'un nouveau format pour le réseau sémantique d'une RSém de lexie, appliqué à la lexie PEIGNE1a**

D'abord, comme mentionné plus haut, nous proposons de diviser systématiquement le réseau en blocs définitionnels. Un bloc est représenté par un encadré entourant les

sémantèmes qui le composent. Chacun de ces blocs est muni d'un en-tête de bloc explicitant son rôle informationnel. Les blocs sont ordonnés du haut vers le bas, selon leur pertinence sémantique pour faciliter l'interprétation de la RSém. Parce que l'ordre ne peut pas être le seul porteur de sens dans une représentation sémantique, nous numérotions chacun des blocs en ordre décroissant de pertinence (1, 2, 3, etc.). Les blocs sont disposés en forme d'escalier, pour faciliter la lecture. Un bloc définitionnel peut contenir des sous-blocs définitionnels, eux aussi organisés de haut en bas selon leur pertinence et numérotés de façon à ce que leur hiérarchisation soit explicite (1.1, 1.2, 1.3, etc.).

À l'intérieur d'un bloc, le nœud communicativement dominant, le ND, se trouve en haut à gauche, encore une fois, pour faciliter la lecture. Des règles directrices devraient gérer la disposition des autres sémantèmes composant le bloc, mais plus de recherche doit être faite à ce sujet. Pour l'instant, nous avons tendance à mettre les sémantèmes qui servent de modificateurs dans les positions inférieures.

### **3.3 *Le rôle du schéma de définition***

Nous croyons que la pertinence sémantique, qui dicte l'ordre de lecture des composantes, est déterminée par un schéma de définition associé au champ sémantique auquel appartient la lexie définie. Ce schéma offre de l'information pour la lecture de chacune des formalisations, mais s'avère absolument nécessaire pour la lecture du réseau sémantique qui ne contient, pour sa part, aucune linéarité. Pour la BDéf et les définitions DEC, le schéma de définition est plutôt utilisé comme une aide lors de leur rédaction. L'ordre de lecture transparait ensuite du format même, c'est-à-dire linéaire, de ces formalisations.

Un schéma de définition est construit pour un champ sémantique donnée. La structure de la définition d'une lexie appartenant à ce champ sémantique s'inspire de ce schéma. Il est possible qu'il y ait plus ou moins de blocs dans la RSém d'une lexie que ce qui est prévu dans le schéma. Par contre, nous ne savons pas s'il est possible que l'ordre des blocs change par rapport à l'ordre prévu par le schéma, c'est-à-dire que la pertinence sémantique d'une composante ne serait pas la même pour une lexie particulière que pour les autres lexies qui partagent son champ sémantique.

Le schéma est adapté au fur et à mesure que des définitions sont écrites. Certaines

composantes s'ajoutent, d'autres deviennent optionnelles. Ainsi, il serait souhaitable de faire une base de données informatisée pour recenser les schémas de définitions qui ont été construits jusqu'à aujourd'hui. Cette base pourrait être modifiée à mesure que les travaux avancent.

### 3.4 Les actants

Les variables ont été extraites du réseau et placées à droite, à l'extérieur des blocs définitionnels. Le fait que les variables se trouvent en dehors des blocs ne devrait pas surprendre. Dans des réseaux divisés de façon similaire, la composante générique est toujours tenue à l'extérieur des blocs même si elle représente souvent le premier actant du sémantème communicativement dominant du sous-réseau. Cela en aide beaucoup la lisibilité puisqu'une même variable peut faire partie de plusieurs blocs définitionnels élémentaires (par ex., la variable 'X' peut se trouver à la fois dans la composante générique et dans les composantes périphériques « Fonction » et « Mode d'utilisation », comme c'est le cas dans la définition de PEIGNE1a (voir Réseau sémantique 4). De plus, une même variable peut se trouver à la fois dans le thème et dans le rhème. Elle peut aussi se trouver dans le posé et dans le présupposé en même temps. C'est le cas dans la définition de SE BERCER1.1.

SE BERCER1.1 (DECFC-IV : 118)

*X se berce dans Y* = Se trouvant sur un artefact Y, || individu X cause de se balancer doucement et régulièrement dans le but de donner à X une sensation agréable.

Dans la définition de SE BERCER 1.1, la partie présupposée 'Se trouvant sur un artefact Y' devrait en fait contenir la variable actancielle 'X'. Elle n'y est pas pour alléger la définition :

SE BERCER1.1

*X se berce dans Y* = X se trouvant sur un artefact Y, || individu X cause de se balancer doucement et régulièrement dans le but de donner à X une sensation agréable.

Ainsi, la variable actancielle 'X' se trouve à la fois dans le posé et dans le présupposé de la définition.

Les arcs qui pointent sur des variables sont constitués d'une ligne pointillée, et non d'une ligne pleine, pour indiquer que les variables dans le réseau ne sont, en fait, qu'un pointeur vers la fiche des variables, qui explicite chacune des variables, ainsi que les contraintes sur les actants et les relations entre les actants. Cette façon de représenter les variables fait

ressortir leur caractère particulier, puisque les variables qui ne sont pas des sémantèmes comme les autres, mais des positions actanciennes en attente d'être remplies par de vrais sémantèmes.

Un bloc définitionnel portant l'en-tête « Caractérisation » contient les caractérisations de variables. Ce bloc définitionnel se trouve à la droite du réseau et la variable caractérisée pointe le bloc en question.

### 3.5 Composante générique

Dans un réseau sémantique, la composante générique sera désormais identifiée par un bloc définitionnel, comme cela est fait dans la BDéf. Comme nous l'avons vu, certaines définitions, notamment des définitions d'ARTEFACTS, exigent la présence de deux composantes génériques. Nous représentons chacune de ces composantes génériques dans le réseau, en prenant soin de numéroter leur en-tête: « Composante générique 1 », « Composante générique 2 », abrégés « CG1 » et « CG2 ». Les deux composantes génériques sont liées par le sémantème 'être' quand le sens de « CG1 », lorsqu'il est décomposé sémantiquement, contient le sens de « CG2 ». Par exemple, dans la définition de PEIGNE1a la « CG1 » 'ustensile' contient la « CG2 » 'objet'.

Les deux composantes génériques sont liées par un 'et/ou' (le 'ou' inclusif) quand le sens de « CG1 » peut être exprimé en même temps que le sens de « CG2 ». L'exemple classique est celui de LIVRE-2 [*Marie lit un très beau/bon livre.*], qui peut désigner autant l'objet que son contenu et, parfois, les deux en même temps. La composante générique serait quelque chose comme 'volume imprimé et/ou son contenu'.

Il y a aussi des cas où deux composantes génériques sont liées par un 'ou' exclusif, quand le sens de « CG1 » ne peut être exprimé en même temps que le sens de « CG2 ». Par exemple la lexie PLOMBIER-2 [*Le plombier doit venir réparer le robinet.*], dont la composante générique est 'individu ou profession' et sera représentée ainsi :

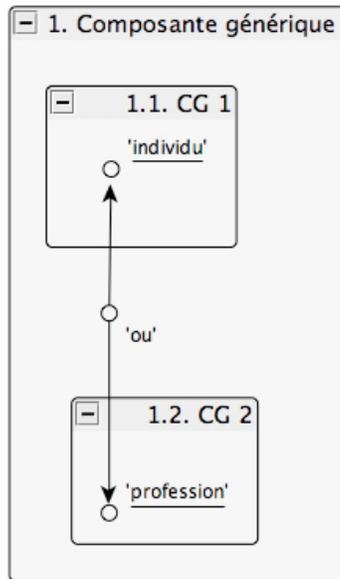


Figure 18 : Composante générique du réseau sémantique de la lexie PLOMBIER-2

On ne peut pas utiliser cette lexie pour dénoter à la fois l'individu et la profession. Par contre, la dérivation individu/profession est régulière et on ne voudrait pas créer deux acceptions séparées pour chacune des professions. Il existe peut-être d'autres cas de figure de lexies à composantes génériques multiples. Ils ne seront pas traités dans le cadre de ce travail.

### 3.6 Les liens de polysémie

Les liens métonymiques et métaphoriques sont indiqués dans le réseau dans l'en-tête de bloc, à côté du rôle informationnel du bloc. Nous croyons qu'il faut aussi expliciter la composante qui indique le type de lien métonymique pour faire un parallèle avec la composante 'comme si' identifiant la métaphore. Ces lexies, ainsi que la composante 'comme si', seront encadrées.

Dans la définition de BERCEAU<sup>II.1</sup> :

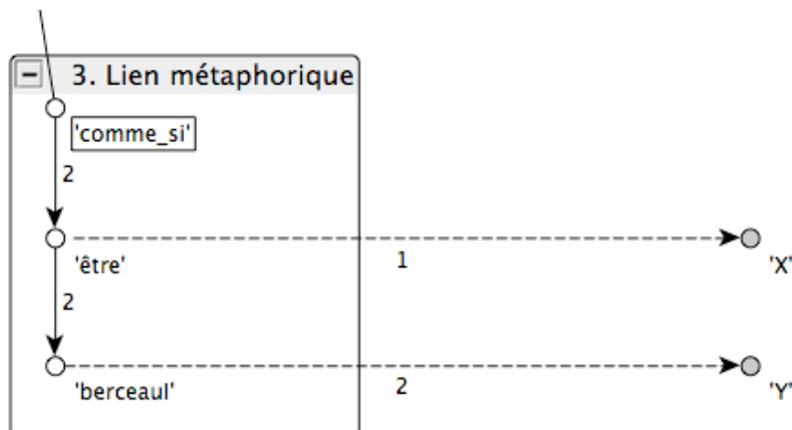


Figure 19 : Illustration d'un en-tête de lien de polysémie

### 3.7 La présupposition

Les blocs qui correspondent à des composantes présupposées sont ombragés. Ceci permet, à la fois, de délimiter la composante au niveau de la SSém en fonction de son rôle informationnel et d'indiquer son statut communicatif.

Notons que la fiche des variables, introduite ci-dessus, est aussi ombragée puisque les données qu'elle contient (contraintes actanciennes et relations sémantiques entre les actants) sont également des éléments présupposés.

### 3.8 La locutionnalité

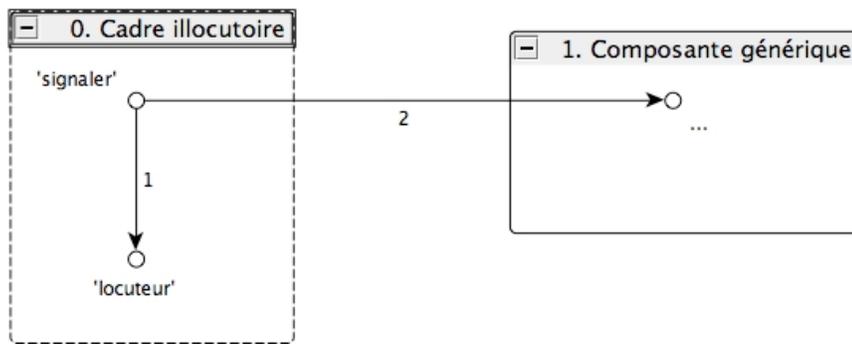
Les sémantèmes qui annoncent le cadre illocutoire d'une lexie sont mis dans un encadré dont le pourtour est pointillé. Par exemple, pour la définition de SÛR<sup>I.2b</sup> :

SÛR<sup>I.2b</sup> (DECFC-IV : 321)

[*Sûr que tu viens demain?*]

**fam.** *Sûr que Y* = Le locuteur signale que le fait Y est sûr<sup>I.2a</sup> pour lui.

Voici la partie du réseau sémantique qui nous intéresse :

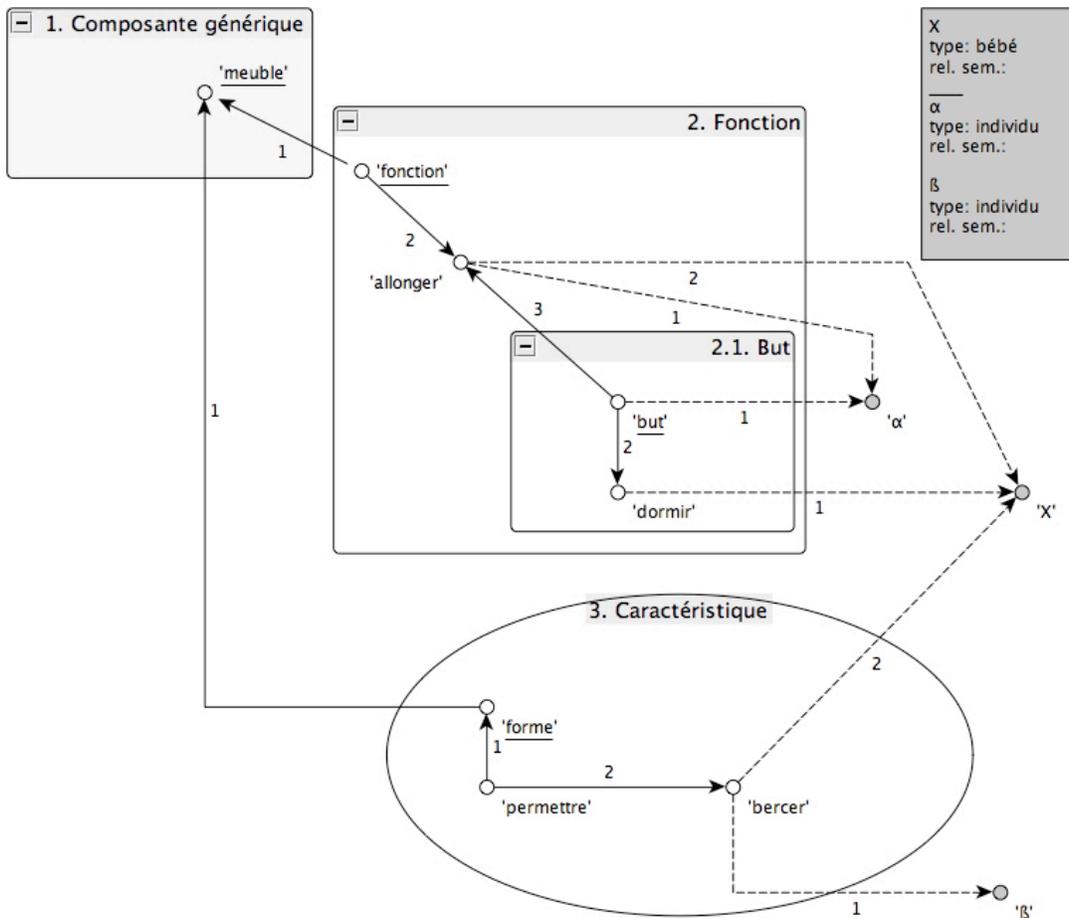


**Figure 20 : Démarcation du cadre illocutoire dans nos réseaux sémantiques**

Notons que, comme pour la présupposition, c'est un encadré qui indique une opposition communicative dans le réseau.

### **3.9 Les composantes faibles et optionnelles**

Les composantes faibles et optionnelles sont encerclées pour marquer leur statut particulier. Elles sont ainsi distinguées des autres sous-réseaux qui, eux, sont encadrés.



**Réseau sémantique 5 : Format proposé appliqué au réseau sémantique de la RSém de la lexie BERCEAU II.1**

Rappelons que 'α' et 'β' représentent des variables non actancielles.

### 3.10 Organisation des blocs et liens entre eux-ci

Idéalement, il faudrait non seulement expliciter les blocs élémentaires dans le réseau sémantique, mais aussi expliciter les liens entre eux. Comme le souligne Wierzbicka (1985 : 53) :

« [...] cette sorte de schéma refléterait non seulement l'ordre des composantes dans l'explication [la définition], mais aussi les liens plus ou moins proches entre elles, qui correspondent aux relations postulées entre les différentes parties du sens. » [Notre traduction. Le soulignement est de Wierzbicka.]

Il nous apparaît que le réseau sémantique serait la meilleure formalisation pour ainsi faire ressortir des liens entre des blocs de définitions qui composent la structure du sens de la lexie.

Par exemple, Wierzbicka est d'avis que, dans les définitions d'ARTEFACT, la forme de l'artefact dépend de la fonction de celui-ci. S'il en est ainsi, ce fait devrait être reflété dans la formalisation des lexies du champ sémantique ARTEFACT, peut-être à l'aide de flèches entre les blocs indiquant le type de lien qui les unit. Ce n'est pas le cas dans nos propositions actuelles. Il s'agit d'un sujet digne d'approfondissement que nous ne pourrions traiter dans le cadre de ce travail.

3.11 Autres exemples de réseaux de lexies

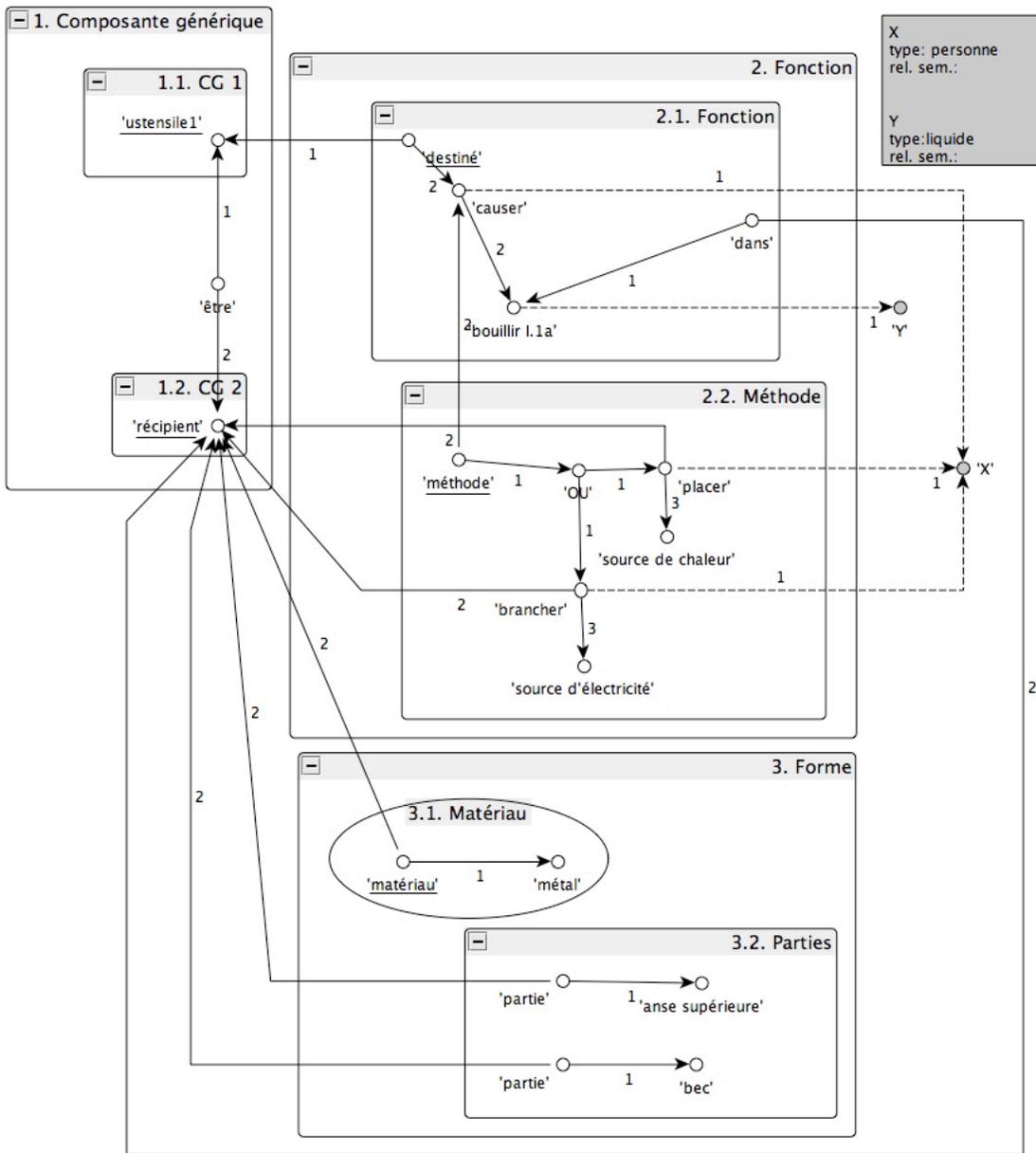


Figure 21 : Format proposé appliqué au réseau sémantique de la RSém de la lexie BOUILLOIRE

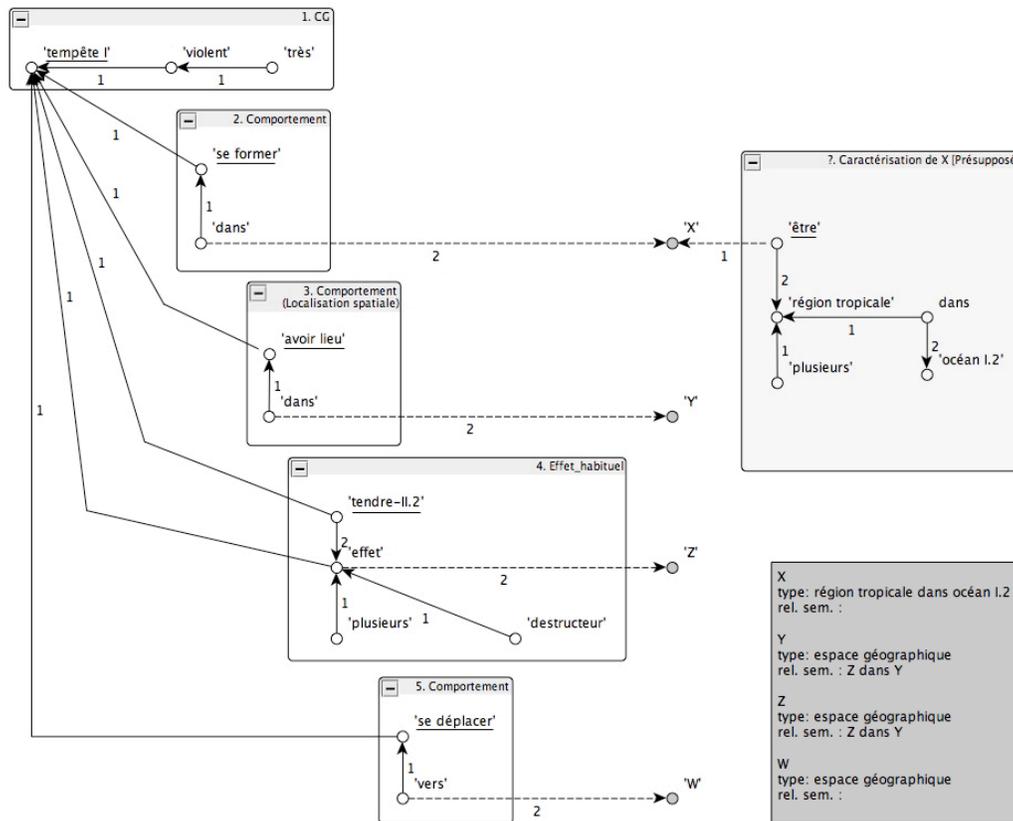


Figure 22 : Format proposé appliqué au réseau sémantique de la RSém de la lexie OURAGANI

Notre proposition n'est pas sans rappeler celles de Dostie *et al.* (1999) et de Polguère (2002a) présentées à la section 2.2.2.1 du chapitre 3. Elle s'en distingue par sa structure de réseau connecté, disposé en escalier, ainsi que par la présence d'une fiche des variables.

#### 4 Modifications à apporter à la définition DEC

Notre proposition s'inspire grandement des formalisations de Wierzbicka et du *Dire Autrement*.

##### 4.1 Démarcation de la composante générique et du NDg

Nous avons vu que ni la composante générique ni le NDg ne sont identifiés dans une définition DEC. Nous croyons que le NDg doit être souligné afin de refléter le formalisme déjà utilisé dans le réseau sémantique. La structuration en composantes génériques et périphériques est au cœur de la structure sémantique des lexies décrites dans un DEC. Aussi, nous considérons trois possibilités pour expliciter cette structure : l'ajout de formalismes typographiques, l'emploi de l'indentation ou les deux en même temps.

Dans L'ILEC, on propose, sans pour autant l'appliquer, l'usage des PETITES MAJUSCULES pour identifier la composante générique dans la définition (Mel'čuk *et al.*, 1995 : 104). Nous favorisons le caractère gras puisque les petites majuscules sont déjà employées dans la TST pour identifier les lexies. Voici des exemples de définitions dans lesquelles le NDg est souligné et la composante générique apparaît en caractères gras :

OURAGAN<sup>I</sup>

*Ouragan de X / de Y [affectant Z et se déplaçant vers W] = Tempête **d'extrême violence** qui se forme dans les régions tropicales X d'un océan<sup>I.2</sup> qui a lieu dans un espace géographique Y qui tend à avoir des effets destructeurs sur Z se trouvant dans Y et qui se déplace vers W.*

BERCEAU<sup>I</sup>

*Berceau de X = Meuble destiné à ce qu'un bébé X y soit allongé pour y dormir (et tel qu'on peut y bercer<sup>I</sup> X).*

L'indentation consiste en l'ajout d'alinéas et de retours de chariot pour séparer les composantes entre elles. Le retour de chariot signifie qu'on est en présence d'un différent type de composante, alors que l'indentation permet de signifier un certain niveau de pertinence sémantique. L'indentation rappelle la façon de faire de Wierzbicka. C'est aussi ce qui est utilisé dans *Dire Autrement*. Une telle technique présente l'avantage d'identifier clairement la séparation entre chacune des composantes périphériques. Voici des exemples de définitions DEC indentées :

OURAGAN<sup>I</sup>

*Ouragan de X / de Y [affectant Z et se déplaçant vers W] =*

**Tempête d'extrême violence**

qui se forme dans les régions tropicales X d'un océan<sup>I.2</sup>

qui a lieu dans un espace géographique Y

qui tend à avoir des effets destructeurs sur Z se trouvant dans Y

et qui se déplace vers W.

PEIGNE<sup>I.a</sup>

*Peigne de X [pour peigner Y de Z] =*

**Ustensile<sup>I</sup> de toilette**

destiné à ce que l'individu X peigne<sup>I</sup> les cheveux ou le poil Y de Z

**objet**

plat

assez rigide

censé être tenu à la main<sup>I.a</sup> par X,

et muni d'une ou de deux rangée(s) de dents<sup>II.1</sup>.

BERCEAU<sup>I</sup>

*Berceau de X =*

**Meuble**

destiné à ce qu'un bébé X y soit allongé

pour y dormir

(et tel qu'on peut y bercer<sup>I</sup> X).

Bien sûr, l'usage de l'indentation et du caractère gras est redondant. Pourtant, nous tenons à conserver les deux pour que la composante générique soit toujours identifiée, même si l'on décide de se passer de l'indentation afin que la définition se rapproche davantage du format de la définition traditionnelle.

#### **4.2 Contraintes sur les actants**

Barque (2008), Milićević (2008), ainsi que d'autres auteurs indiquent les contraintes sur les actants en les mettant en indices rattachés à la variable caractérisée. Certains positionnent l'indice avant la variable, pour suivre l'ordre de la langue naturelle, d'autres, le mettent après. Nous choisissons de le placer avant pour faciliter la lecture. Aussi, pour éviter la redondance, nous n'indiquons la contrainte qu'au premier usage de la variable. Voici des exemples d'explicitation des contraintes actanciennes :

**OURAGAN<sup>I</sup>**

*Ouragan de X / de Y [affectant Z et se déplaçant vers W] =*

**Tempête<sup>I</sup> d'extrême violence**

qui se forme dans les régions tropicales X d'un océan<sup>I.2</sup>

qui a lieu dans un <sub>espace géographique</sub> Y

qui tend à avoir des effets destructeurs sur <sub>espace géographique</sub> Z se trouvant dans Y

et qui se déplace vers <sub>espace géographique</sub> W.

**BERCEAU<sup>I</sup>**

*Berceau de X =*

**Meuble**

destiné à ce qu'un <sub>bébé</sub> X y soit allongé

pour y dormir

(et tel qu'on peut y bercer<sup>I</sup> X).

**4.3 Les connecteurs**

Nous avons vu que les définitions BDéf contiennent des connecteurs sémantiquement vides. Il serait possible de simplement enlever ces connecteurs, mais nous croyons qu'ils facilitent la lecture de la définition en langage naturel, surtout si l'on choisit de présenter la définition « à plat », c'est-à-dire sans indentation. Ainsi, nous utiliserons plutôt une couleur de police grise en gras pour indiquer qu'il ne s'agit pas d'un vrai sémantème. Voici des exemples d'identification des opérateurs sémantiquement vides :

**PEIGNE<sup>1a</sup>**

[...]

**Ustensile<sup>1</sup> de toilette**

[...]

plat

assez rigide

censé être tenu à la main<sup>I.a</sup> par <sub>individu</sub> X,

et muni d'une ou de deux rangée(s) de dents<sup>II.1</sup>.

PEIGNE**Ia** :

*Peigne de X [pour peigner Y de Z] = Ustensile**I** de toilette destiné à ce que <sub>individu</sub>X peigne**I** les <sub>cheveux ou le poil</sub>Y de Z – objet plat assez rigide censé être tenu à la main**I.a** par X, et muni d’une OU de deux rangée(s) de dents**II.1**.*

#### 4.4 Organisation en blocs définitionnels

L’indentation semble être la meilleure façon d’identifier et d’étiqueter chacune des composantes périphériques en fonction de son rôle informationnel. Les en-têtes se trouvent à gauche de la définition. À chaque composante périphérique correspond un en-tête.

PEIGNE <b>Ia</b>	
<i>Peigne de X [pour peigner Y de Z] =</i>	
CG 1 :	<b>Ustensile<b>I</b> de toilette</b>
Fonction :	destiné à ce que <sub>individu</sub> X peigne <b>I</b> les <sub>cheveux ou le poil</sub> Y de Z
CG 2 :	<b>objet</b>
Forme :	plat
Matériel :	assez rigide
Mode d’utilisation :	censé être tenu à la main <b>I.a</b> par X,
Parties :	et muni d’une ou de deux rangée(s) de dents <b>II.1</b> .

#### 4.5 Les liens de polysémie

Une fois la définition DEC indentée et chacune de ses composantes spécifiques munie d'un en-tête, il est aisé d'apposer un en-tête précis spécifiant le type de lien de polysémie qui est en jeu dans une définition (lien métaphorique ou métonymique). Comme dans le réseau sémantique, nous encadrons les composantes qui introduisent le pont sémantique. Voici un exemple avec BERCEAU**II.1** [*La Nouvelle-Orléans, berceau du blues, est en deuil*] :

BERCEAU <b>II.1</b> <i>Berceau [X] de Y =</i>	
<b>Composante générique</b>	<u>Lieu X</u>
<b>Caractérisation de X</b>	où a commencé l'existence de Y
<b>[lien métaphorique]</b>	[ <u>comme si</u> X était le berceau <b>I</b> de Y].

Aussi, pour BOL**b**, qui est en relation de métonymie avec BOL**a**, ci-dessous :

BOL <b>a</b> (DECFC-IV : 138) <i>[Il boit son café dans un bol]</i> Bol de X à Y = Pièce de vaisselle destinée à contenir une boisson <b>I</b> ou un met Y qui sera bu <b>I.1</b> ou mangé <b>I.1a</b> par un être X – récipient de forme hémisphérique de la grandeur de la main <b>la</b> .
---

**Définition 18: Définition DEC de BOL**a****

BOL <b>b</b> (DECFC-IV : 138) <i>[Il boit un bol de café au lait]</i> Bol de Y =	
<b>Composante générique</b>	<u>Quantité de Y</u>
<b>[lien métonymique]</b>	qu'un bola <u>peut contenir</u> .

#### 4.6 La locutionnalité

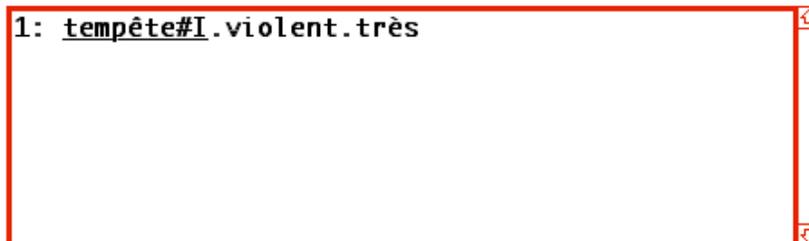
Nous avons vu dans la section précédente la manière dont nous indiquons les sémantèmes qui ont trait au cadre illocutoire dans un réseau sémantique. Le même formalisme est utilisé dans la définition DEC – le cadre illocutoire est mis dans un encadré pointillé :

**SÛRI.2b**[- *Sûr que je viens!*]**fam.** *Sûr que Y* = [Le locuteur signale que] le fait Y est sûr**I.2a** pour lui.**4.7 Remarque sur la définition DEC**

Certains pourraient croire que nous avons alourdi le format des définitions DEC, des définitions qui étaient déjà difficiles à lire pour les non-initiés. On peut imaginer que ces définitions pourraient être utilisées dans un dictionnaire informatisé, dans lequel il serait possible de faire apparaître ou disparaître certaines informations, par exemple les contraintes sur les actants ou les en-têtes de blocs. Nous nous inspirons pour cela du *Dire Autrement*, qui permet aux apprenants de consulter des définitions qui ne sont pas trop lourdes de formalismes tout en leur donnant accès à toute l'information, s'ils désirent en voir plus. Le DAFLES (Selva *et al.*, 2002), un dictionnaire en ligne pour les apprenants du français, offre ainsi la possibilité de consulter ses définitions dans leur version longue ou courte.

**5 Modifications à apporter à la définition BDéf****5.1 Démarcation de la composante générique et du nœud communicativement dominant**

Le NDg doit être souligné afin d'uniformiser les formalismes utilisés dans les formalisations, en voici un exemple :



1: tempête#I.violent.très

Figure 23 : Bloc central de la définition BDéf de OURAGANI

**5.2 Organisation en blocs définitionnels**

Bien sûr, la BDéf possède déjà une séparation en blocs. Pourtant, nous remarquons que ces blocs se trouvent tous sur le même niveau hiérarchique. Lorsque l'on regarde les définitions indentées de Wierzbicka ou du *Dire Autrement*, la nécessité d'utiliser une hiérarchisation

des blocs, en blocs et en sous-blocs, se fait sentir.

Il existe déjà un niveau secondaire de numérotation des composantes, mais celui n'est utilisé qu'avec les opérateurs. Il faudrait aussi s'en servir pour indiquer que des composantes sont enchâssées dans d'autres composantes.

### 5.3 *Explicitation de l'ordre de lecture des blocs définitionnels*

Nous avons fait remarquer à la section 2.2.2 du chapitre 3 que rien n'explique l'ordre des blocs définitionnels dans une définition BDéf. Nous croyons que c'est la présence d'un schéma de définition au niveau du champ sémantique qui justifie l'ordre des composantes sémantiques dans une définition.

### 5.4 *Composantes faibles et optionnelles*

Nous traiterons les composantes faibles comme les composantes optionnelles, c'est-à-dire en les mettant entre parenthèses. Le manque de données concernant ce phénomène sémantique ne nous permet pas de prévoir un formalisme plus précis.

### 5.5 *Les liens de polysémie*

Pour expliciter la composante qui indique le type de lien métonymique, nous utilisons l'encadré. Nous faisons de même pour la métaphore.

```

: /*Lien_métaphorique*/
3: *2 _comme_si_ *4
4: *1 être berceau#I de Y

```

### 5.6 *La présupposition*

Rappelons que, dans la mesure du possible, les éléments appartenant à des structures distinctes (SSém et SSém-Comm) sont marqués de façons différentes pour mettre en évidence leur appartenance respective. Par exemple, dans la BDéf, les en-têtes indiquant les composantes présupposées et les composantes optionnelles sont entre crochets à côté de l'en-tête de bloc correspondant :

```
/*Caractérisation [présupposé]*/
```

```
/*Caractérisation [optionnel]*/
```

Pourtant, le caractère optionnel de la composante appartient à la SSém tandis que son caractère présupposé appartient plutôt à la SSém-Comm. Nous proposons de mettre l'en-



Mel'čuk & Wanner (2001 : 15) :

1. Composante générique
2. Posé
  - 1.1 Action centrale
  - 1.2 Acte de langage
  - 1.3 But (de l'acte de langage)
  - 1.4 Caractéristiques (de l'acte de langage)
3. Présuppositions
4. Actants

Ainsi, nous proposons ce schéma dans lequel toute l'information donnée relève de la SSém et où ce qui a trait à la SSém-Comm se superpose au contenu de ce qui a trait à la SSém. Nous n'indiquons pas le posé, qui est la valeur par défaut. Nos modifications sont soulignées :

1. Composante générique
2. Composantes périphériques
  - 1.1 Action centrale
  - 1.2 Acte de langage
  - 1.3 But (de l'acte de langage)
  - 1.4 Caractéristiques (de l'acte de langage)
  - 1.5 *En-tête de rôle informationnel quelconque* |[présupposé]|
  - etc.
3. Actants

Aussi, comme on le voit dans la composante 1.5, nous sommes d'avis que les composantes présupposées possèdent également un rôle informationnel, c'est-à-dire qu'une composante présupposée devrait également être dotée d'un en-tête indiquant le sémantisme que cette composante apporte à la définition (« Caractérisation », « Matériel », etc.).

## 7 Tableau final

Nous reprenons ici le tableau de la section 4 du chapitre 3. Les nouveaux éléments de formalisme sont indiqués par l'encadrement de la valeur « OUI » ou « NON ». Les cases grises marquent les disparités qui demeurent.

	<b>Réseau sémantique</b>	<b>Définition lexicographique</b>	<b>Définition BDéf</b>
Éléments de formalisme SSém			
Identification de la composante générique	<u>OUI</u> bloc définitionnel	<u>OUI</u> caractères gras et indentation	OUI bloc central
Identification du NDg	OUI <u>soulignement</u>	<u>OUI</u> <u>soulignement</u>	<u>OUI</u> <u>soulignement</u>
Présence des blocs définitionnels	<u>OUI</u> encadrés	<u>OUI</u> indentation	OUI blocs définitionnels
Typages de blocs définitionnels	<u>OUI</u> en-têtes	<u>OUI</u> en-têtes	OUI en-têtes /*En-tête sémantique*/
Ordre des composantes / des blocs	<u>OUI</u> schémas de déf et numérotation	<u>OUI</u> schéma de définition et ordre linéaire langagier	<u>OUI</u> schéma de définition et numérotation : *1, *2, etc.
Liens entre les blocs <sup>9</sup>	NON	NON	NON
Explicitation des liens métaphoriques	<u>OUI</u> en-tête et sémantème 'comme si'	<u>OUI</u> en-tête et sémantème 'comme si'	OUI en-tête et sémantème 'comme si' /*[[lien_métaphorique]*/
Explicitation des liens métonymiques	<u>OUI</u> en-tête	<u>OUI</u> en-tête	OUI en-tête /*[[lien_métonymique]*/
Explicitation des relations sémantiques entre les actants	<u>OUI</u> Fiche des variables	NON	OUI bloc des relations sémantiques
Composantes optionnelles	<u>OUI</u> composante encerclée	OUI (composante opt.)	OUI (composante opt.)
Composantes faibles	<u>OUI</u> composante encerclée	OUI (composante faible) <sub>condition</sub>	OUI (composante faible)
Explicitation du typage d'actant	<u>OUI</u> Fiche sur les d'actant	<u>OUI</u> indice X	OUI Bloc de typage des actants

<sup>9</sup> Nous ajoutons ce concept lexicographique pour signaler que l'ajout de formalismes pour le représenter serait souhaitable.

	<b>Réseau sémantique</b>	<b>Définition lexicographique</b>	<b>Définition BDéf</b>
Distinction/absence des connecteurs sémantiquement vides	OUI Ils sont absents.	<b>OUI</b> caractères gras et couleur de la police grise	OUI Ils sont absents.
Éléments de formalisme SSém-Comm			
Présupposés	OUI sous-réseau ombragé	<b>OUI</b>  [présupposé]	<b>OUI</b> /* [présupposé] */
Locutionnalité	<b>OUI</b> rectangle pointillé	<b>OUI</b> rectangle pointillé	<b>OUI</b> bloc pointillé dans bloc central
Thématicité	NON	NON	NON

**Tableau II:** Tableau résumant les éléments de formalisme proposés pour les concepts lexicologiques analysés dans les trois formalisations traitées

## CHAPITRE 5 Conclusion

Ce mémoire portait sur trois formalisations de la RSém de lexies dans la théorie Sens-Texte : le réseau sémantique, la définition DEC et la définition BDéf. Après avoir décrit ces trois formalisations, nous avons recensé les disparités qui existaient entre ces dernières, pour finalement proposer des modifications visant à pallier les disparités trouvées, soit en ajoutant des éléments de formalisme, soit en uniformisant les formalismes utilisés.

Notre objectif était de rendre les formalismes utilisés dans les formalisations plus uniformes pour en simplifier la lecture, mais aussi pour faciliter la « traduction » d'une formalisation à l'autre, et ce, de façon automatique ou non. L'objectif consistait également à rendre plus complète la représentation sous forme de réseau sémantique, la formalisation idéale selon la LEC. Pour ce faire, nous sommes allés chercher des éléments de formalismes de différentes formalisations existant dans la LEC et ailleurs.

Dans les tableaux 1 et 2, nous avons synthétisé une partie des problématiques à considérer lors de l'élaboration d'une formalisation de la représentation sémantique dans l'approche Sens-Texte. Ces tableaux peuvent servir de point de départ à une réflexion plus approfondie sur les différentes formalisations.

Il importe maintenant de définir un plus grand nombre de lexies en se servant des formalismes proposés, chaque formalisation de lexie étant propice à la découverte de nouveaux problèmes lexicographiques, devant à leur tour être intégrés dans la théorie et formalisés dans les représentations. Aussi, plusieurs problèmes de description ont été soulevés sans que nous ayons pu offrir de solutions dans le cadre de ce travail. On devrait, entre autres, étendre l'étude des différentes formalisations aux RSém d'énoncés. De plus, la représentation des schémas de définitions mériterait une plus grande attention. Nous croyons qu'il faudrait en venir à la création d'une base de données regroupant les schémas de définitions de tous les champs sémantiques décrits dans les travaux de la LEC, en uniformisant leur structure.

## Bibliographie

- Alonso Ramos, M. (2003). Hacia un Diccionario de Colocaciones del español y su codificación. Dans M. A. Martí *et al* (dir.), *Lexicografía computacional y semántica* (pp. 11-34). Barcelone: Universitat de Barcelona.
- Altman, J., & Polguère, A. (2003). La Bdéf: base de définitions dérivée du dictionnaire explicatif et combinatoire. Dans *Actes de la première conférence internationale de théorie Sens-Texte (MTT2003)* (pp. 43-54). Paris.
- Apresjan, Jurij (1974) *Leksičeskaja semantika. Sinonimičeskie sredstva jazyka* [Lexical Semantics. Synonymic Means of the Language]. Moscow: Nauka. Paris.
- Barque, L. (2003). *Opérations sémantiques sur une base de données de définitions Sens-Texte* (Mémoire de DEA). Université Paris 7, Paris.
- Barque, L. (2008). *Description et formalisation de la polysémie régulière du français* (Thèse de doctorat). Université Paris 7, Paris.
- Barque, L., & Nasr, A. (2005). Computing semantic relations on structured lexical definitions. Dans *Proceedings of the Second International Conference on the Meaning Text Theory* (pp. 41-49). Moscou.
- Barque, L., & Polguère, A. (2004). A Definitional Metalanguage for Explanatory Combinatorial Lexicography. Présentation non-publiée faite à l'*International Colloquium on Word Structure and Lexical Systems: models and applications*, Pavia (Italy), 16-17 décembre 2004.
- Dostie, G., Mel'čuk, I., & Polguère, A. (1999). Méthodologie d'élaboration des articles du "Dictionnaire explicatif et combinatoire du français contemporain". Dans Mel'čuk *et al.*, *Dictionnaire explicatif et combinatoire du français contemporain-IV* (pp. 11-27). Montréal: Presses de l'Université de Montréal.

- Fox, C. B. (2007). *Longman dictionary of contemporary english*. Melbourne: Addison Wesley.
- Goddard, C. UNE - Linguistics - NSM Homepage. *NSM Homepage*. Repéré le 11 janvier, 2010, à <http://www.une.edu.au/bcss/linguistics/nsm/>
- Goddard, C., & Wierzbicka, A. (1994). *Semantic and lexical universals: theory and empirical findings*. Amsterdam; Philadelphia: J. Benjamins.
- Gougenheim, G. (1958). *Dictionnaire fondamental de la langue française* (Nouvelle édition revue et augmentée.). Paris: Didier.
- Green, G. (1969). On the notion 'Related Lexical Entry'. *Papers from the Fifth Regional Meeting, Chicago Linguistic Society*, 76-88.
- Hamel, M., & Milićević, J. (2007). Analyse d'erreurs lexicales d'apprenants du FLS : démarche empirique pour l'élaboration d'un dictionnaire d'apprentissage. *Journal of the Canadian Association for Applied Linguistics*, 25-45.
- Iordanskaja, L., & Mel'čuk, I. (1990). Semantics of Two Emotion Verbs in Russian: BOJAT'SJA 'to be afraid' and NADEJAT'SJA 'to hope'. *Australian Journal of Linguistics*, 10, 307-357.
- Lew, R., & Dziemianko, A. (2006). A new type of folk-inspired definition in English monolingual learners' dictionaries and its usefulness for conveying syntactic information. *International Journal of Lexicography*, 19(3), 225-242.
- Mel'čuk, I. (1984). *Dictionnaire explicatif et combinatoire du français contemporain-I*. Montréal: Presses de l'Université de Montréal.
- Mel'čuk, I. (1988a). *Dictionnaire explicatif et combinatoire du français contemporain-II*. Montréal: Presses de l'Université de Montréal.
- Mel'čuk, I. (1988b). *Dependency syntax: theory and practice*. Albany, NY.: State University

of New York Press.

Mel'čuk, I. (1992). *Dictionnaire explicatif et combinatoire du français contemporain-III*.

Montréal: Presses de l'Université de Montréal.

Mel'čuk, I. (1997). *Vers une linguistique Sens-Texte*. Paris: Collège de France.

Mel'čuk, I. (2001). *Communicative organization in natural language: the semantic-communicative structure of sentences*. Amsterdam, Philadelphia: J. Benjamins.

Mel'čuk, I., Arbachevsky-Jumarie, N., Iordanskaja, L., Mantha, S., & Polguère, A. (1999).

*Dictionnaire explicatif et combinatoire du français contemporain-IV*. Montréal: Presses de l'Université de Montréal.

Mel'čuk, I., Clas, A., & Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve: Duculot.

Mel'čuk, I., & Polguère, A. (2007). *Lexique actif du français: l'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*.

Bruxelles: De Boeck.

Mel'čuk, I., & Wanner, L. (2001). Towards a lexicographic approach to lexical transfer in machine translation (Illustrated by the German-Russian language Pair). *Machine Translation*, 16(1), 21-87.

Milićević, J., & Hamel, M. (2007). Un dictionnaire de reformulation pour les apprenants du français langue seconde. Dans G. Chevalier, K. Gauvin & D. Merkle (dir.), *PAMAPLA 29/Actes du 29e Colloque annuel de l'ALPA, Université de Moncton, 4-5 novembre, 2005*. (pp. 145-167).

Milićević, J. (1997). *Étiquettes sémantiques dans un dictionnaire formalisé du type Dictionnaire Explicatif et Combinatoire* (mémoire de maîtrise). Université de Montréal, Montréal.

- Milićević, J. (2007). Semantic Equivalence Rules in Meaning-Text Paraphrasing. Dans L. Wanner (dir.), *Selected lexical and grammatical issues in the meaning-text theory: in honour of Igor Mel'čuk* (pp. 267-296). Amsterdam ; Philadelphia: John Benjamins Pub. Co.
- Milićević, J. (2008). Structure de la définition lexicographique dans un dictionnaire d'apprentissage explicatif et combinatoire. Dans E. Bernal & J. DeCesaris (dir.), *Proceedings of the XIII EURALEX International Congress. Barcelona* (pp. 551-561). Barcelona: University Institute for Applied Linguistics, Pompeu Fabra University.
- Milićević, J. (2009). C'est la définition de quel mot? Tester la validité des définitions lexicographiques pour un dictionnaire d'apprentissage. Dans D. Beck, K. Gerdes, J. Milićević & A. Polguère (dir.), *Proceedings of the Fourth International Conference on Meaning-Text Theory, MTT'09* (pp. 275-285). Université de Montréal.
- Milićević, J., & Polguère, A. (2010). Ambivalence sémantique des noms de communication langagière du français. Dans F. Neveu, V. Muni Toke, J. Durand, T. Klingler, L. Mondada, S. Prévost (dir.), *Congrès Mondial de Linguistique Française - CMLF 2010* (pp. 1029-1050), Paris: Institut de Linguistique Française.
- Padučeva, E. (1997). Verb Categorization and the Format of a Lexicographic Definition (Semantic Types of Causative Relations). Dans L. Wanner (dir.), *Recent trends in meaning-text theory* (pp. 61-72). Amsterdam; Philadelphia: John Benjamins Pub. Co.
- Padučeva, E. (2009). Event structure: Taxonomy, semantic roles, aspect, causation. *Automatic Documentation and Mathematical Linguistics*, 43(3), 196-202.
- Padučeva, E., & Kustova. (2004). Lexicographer. *RUSLING*. Repéré le 4 mars, 2010 à

<http://www.rusling.narod.ru/>

Polguère, A. (1990). *Structuration et mise en jeu procédurale d'un modèle linguistique déclaratif dans un cadre de génération de texte* (Thèse de doctorat). Université de Montréal, Montréal.

Polguère, A. (1992). Remarques sur les réseaux sémantiques Sens-Texte. Dans A. Clas (dir.), *Le Mot, les mots, les bons mots : hommage à Igor A. Mel'čuk par ses amis, collègues et élèves à l'occasion de son soixantième anniversaire = Word, words, witty words* (pp. 109-148). Montréal: Presses de l'Université de Montréal.

Polguère, A. (1998). La théorie sens-texte. *Dialangue*, 8, 9-30.

Polguère, A. (2002a). Le sens linguistique peut-il être visualisé? Dans D. Lagorgette & P. Larrivée (dir.), *Représentations du sens linguistique* (pp. 89–103), Volume 22 de *Lincom Studies in Theoretical Linguistics*. Munich: Lincom Europa.

Polguère, A. (2002b). Modélisation des liens lexicaux au moyen des fonctions lexicales. *TALN 2002*, 37-60.

Polguère, A. (2003a). *Lexicologie et sémantique lexicale : notions fondamentales*. Paramètres. Montréal: Presses de l'Université de Montréal.

Polguère, A. (2003b). *Étiquetage sémantique des lexies dans la base de données DiCo*. *TAL*, 44(2), 39-68.

Polguère, A. (2009). *Hiérarchie des étiquettes sémantiques*. Repéré le 7 avril, 2010, à <http://olst.ling.umontreal.ca/dicopop-laf/hierarchie.php>

Pustejovski, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.

Rey, A. (1977). *Le lexique : images et modèles, du dictionnaire à la lexicologie*. Paris: Armand Colin.

Rey, A., P. (2005). *Le grand Robert de la langue française [ressource électronique]*. (2<sup>e</sup>

- éd.). Paris: Dictionnaires Le Robert.
- Rey-Debove, J., & Rey, A. (2010). *Le nouveau petit Robert : dictionnaire alphabétique et analogique de la langue française* (Nouv. éd. du Petit Robert.). Paris: Dictionnaires Le Robert.
- Rundell, M. (1998). Recent trends in English pedagogical lexicography. *International Journal of Lexicography*, 11(4), 315.
- Selva, T., Verlinde, S., & Binon, J. (2002). Le DAFLES, un nouveau dictionnaire électronique pour apprenants du français. Dans A. Braasch & C. Povlsen (dir.), *Proceedings of the Tenth International Euralex Congress, Euralex 2002* (pp. 773–83), Copenhagen: Center for Sprogteknologi.
- Wanner, L., Alonso Ramos, M., & Martí, A. (2004). Enriching the Spanish EuroWordNet by collocations. Dans *Proceedings of LREC 2004*, Paris: ELRA/ELDA.
- Wierzbicka, A. (1972). *Semantic primitives*. Francfort/M.: Athenäum-Verl.
- Wierzbicka, A. (1977). Mental Language and Semantic Primitives. *Communication and Cognition Ghent*, 10(3-4), 155-179.
- Wierzbicka, A. (1980). *Lingua mentalis : the semantics of natural language*. Sydney; New York: Academic Press.
- Wierzbicka, A. (1985). *Lexicography and conceptual analysis*. Ann Arbor: Karoma.
- Wierzbicka, A. (1987). *English speech act verbs: a semantic dictionary*. Sydney; Orlando: Academic Press.
- Wierzbicka, A. (1996). *Semantics: primes and universals*. Oxford [England]; Toronto: Oxford University Press.