

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Vers une grammaire d'unification Sens-Texte du français:
le temps verbal dans l'interface sémantique-syntaxe

par
François Lareau

Thèse de doctorat effectuée en cotutelle au

Département de linguistique et de traduction
Faculté des arts et des sciences
Université de Montréal

et au

Département de linguistique
Université Paris Diderot (Paris 7)



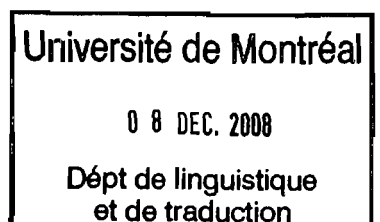
Thèse présentée à la Faculté des études supérieures de
l'Université de Montréal en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.) en linguistique.

et à

l'Université Paris Diderot (Paris 7) en vue de l'obtention du
grade de Docteur en linguistique

Août 2008

© François Lareau, 2008



Page d'identification du jury

Université de Montréal
Faculté des études supérieures

et

Département de linguistique
Université Paris Diderot (Paris 7)

Cette thèse intitulée
*Vers une grammaire d'unification Sens-Texte du français : le
temps verbal dans l'interface sémantique-syntaxe*

présentée et soutenue à l'Université de Montréal par
François Lareau

a été évaluée par un jury composé des personnes suivantes :

Président-rapporteur et membre du jury	Dr. Alain Polguère
Directeur de recherche (Université de Montréal)	Dr. Igor Mel'čuk
Directeur de recherche (Université Paris Diderot)	Dr. Sylvain Kahane
Membre du jury	Dr. Laurence Danlos
Examineur externe	Dr. Owen Rambow
Représentant du doyen de la FES	Dr. Antoine Soare

Résumé

Cette thèse vise deux objectifs principaux. D'une part, nous cherchons à décrire la flexion verbale en français, et plus particulièrement celle en temps dans l'interface sémantique-syntaxe. D'autre part, nous voulons, à travers ce phénomène linguistique central, développer le formalisme de la grammaire d'unification Sens-Texte (GUST) et en vérifier l'utilité pour la description des langues.

Pour atteindre notre premier objectif, nous proposons une méthodologie, inspirée de la lexicographie, pour l'étude des signes grammaticaux. Cette méthodologie met l'accent sur les signes dans leur ensemble, et non sur une seule de leurs composantes (signifié, signifiant ou combinatoire).

Au niveau descriptif, notre thèse principale est que le français possède non pas une, mais bien deux catégories flexionnelles de temps complémentaires :

- 1) La catégorie de décalage, qui regroupe deux grammèmes situant un point de repère par rapport au moment d'énonciation :
 - Le grammème **non-décalé** signifie que le point de repère est 'maintenant' ou un fait dans le futur. Son signifiant est nul.
 - Le grammème **décalé** indique que le repère est un fait dans le passé. Il s'exprime par le suffixe **-AI-**.
- 2) La catégorie de temps comme tel, qui regroupe trois grammèmes situant les faits par rapport à ce point de repère :
 - Le grammème **simultané** signifie que le fait dénoté par le verbe est simultanément à son point de repère. Son signifiant est nul.
 - Le grammème profond **antérieur** indique que le fait est antérieur à son point de repère. Son signifiant est l'auxiliaire AVOIR (ou ÊTRE). Dans un registre littéraire, il s'exprime conjointement avec le grammème **non-décalé** par le suffixe du passé simple.
 - Le grammème **postérieur** pose le fait comme postérieur à son point de repère. Il s'exprime par le suffixe **-R-**.

Au niveau formel, nos principaux apports au formalisme de GUST sont la représentation des structures en termes d'objets et de fonctions, ainsi que la modélisation des décompositions sémantiques. La formalisation de notre modèle de la flexion verbale s'avère un moyen intéressant de vérifier l'utilité de GUST puisqu'elle met en jeu des signes variés. Elle nous donne également l'occasion d'observer le mécanisme d'articulation des modules de la grammaire.

Mots-clés: Temps grammatical; Grammaire d'unification Sens-Texte; Interface sémantique-syntaxe; Flexion; Verbes; Français (langue)

Abstract

This thesis, entitled *Towards a Meaning-Text unification grammar of French: verbal tense in the semantics-syntax interface*, has two main objectives. First, we aim at describing French verbal inflection, more particularly tense in the semantics-syntax interface. Second, we want to develop the Meaning-Text unification grammar (MTUG) formalism and test its adequacy for language description.

To achieve our first goal, we propose a lexicography-inspired methodology for the study of grammatical signs. This methodology is centered on the linguistic sign as a whole, not only on one of its components (signified, signifier or syntactics).

At the descriptive level, our main thesis is that French has not only one, but two inflectional categories of tense that complement each other:

- 1) The category of shifting (in French, *décalage*), which comprises two grammemes that locate a temporal reference point in relation with the time of speech:
 - The grammeme **non-shifted** indicates that the reference point is either ‘now’ or a future point in time. Its signifier is null.
 - The grammeme **shifted** indicates that the reference point is in the past. It is expressed by the suffix –AI–.
- 2) The category of tense proper, which comprises three grammemes that situate facts in relation with this reference point:
 - The grammeme **simultaneous** means that the fact denoted by the verb is simultaneous to the reference point. Its signifier is null.
 - The deep grammeme **anterior** indicates that the fact is before the reference point. It is expressed by the auxiliary AVOIR (or ÊTRE). In literary speech, it is expressed jointly with the grammeme **non-shifted** by the simple past suffix.
 - The grammeme **posterior** indicates that the fact is after the reference point. It is expressed by the suffix –R–.

At the formal level, our main contributions to the MTUG formalism are the representation of structures in terms of objects and functions, as well as the modeling of semantic decompositions. The formalization of our French inflection model turns out to be an interesting way of testing MTUG's adequacy, since this phenomenon involves signs of different natures. It also provides a good opportunity to observe the articulation of the grammar's modules.

Keywords: Tense; Meaning-Text unification grammar; Semantics-syntax interface; Inflection; Verbs; French (language)

Table des matières

Résumé	iii
Abstract	v
Table des matières	vii
Liste des tableaux	xii
Liste des figures	xiii
Conventions typographiques	xvii
Symboles utilisés	xvii
Liste des abréviations	xviii
Remerciements	xx
I. Introduction	21
<hr/>	
1. Problématique de la recherche.	21
2. Pourquoi le temps grammatical?	23
3. Organisation de la thèse.	24
II. Délimitation du champ d'étude	25
<hr/>	
1. Précision terminologique.	26
2. Les formes synthétiques	27
3. Les formes analytiques à un auxiliaire	28
4. Les formes analytiques à deux auxiliaires	31
5. Les formes analytiques à trois auxiliaires	33
III. La théorie Sens-Texte	35
<hr/>	
1. Présentation générale.	35
2. Les niveaux de représentation	38
2.1. La représentation sémantique	38
2.2. La représentation syntaxique profonde	40
2.3. La représentation syntaxique de surface	42
2.4. Les représentations morphologiques et phonologiques	42
3. Les modules de correspondance	43
4. Synthèse.	45

IV. L'étude du grammème : théorie et méthodologie **46**

1. Les notions de grammème et de grammie	46
2. Flexion vs dérivation: propriétés distinctives	51
3. Une approche discrète en deux étapes	52
4. La grammie de base d'un grammème	53
4.1. Les critères basés sur le sens	53
4.2. Les critères basés sur la combinatoire	57
5. La description du sens grammatical	60
6. Le regroupement des grammèmes en catégories flexionnelles	62
6.1. Les critères basés sur la combinatoire	63
6.2. Les critères basés sur le sens	65
7. Grammème et phraséologie	67
8. Synthèse	68

V. Revue critique de quelques modèles de la conjugaison française **70**

1. Le modèle traditionnel	72
2. Imbs	80
3. Les fonctionnalistes	84
3.1. Martinet	84
3.2. Touratier	91
4. Wilmet	96
5. Gosselin	106
6. Synthèse	113

VI. Un modèle descriptif de la conjugaison française **115**

1. La finitude	116
1.1. L'infinitif	118
1.2. Les participes	121
1.2.1. <i>Le participe présent</i>	123
1.2.2. <i>Le participe passé</i>	124
1.3. Le gérondif	129
1.4. Les verbes finis	130
1.5. Synthèse	131
2. La voix	132
3. Le mode	134
3.1. L'indicatif	135

3.2. Le subjonctif	137
3.3. L'impératif	141
3.4. Le conditionnel	142
3.5. Synthèse	144
4. Les formes surcomposées	144
4.1. Définition et acceptabilité des formes surcomposées	145
4.1.1. <i>Le passé surcomposé</i>	146
4.1.2. <i>Le plus-que-parfait surcomposé</i>	150
4.1.3. <i>Le conditionnel surcomposé</i>	150
4.1.4. <i>Le futur antérieur surcomposé</i>	152
4.1.5. <i>Les autres «formes surcomposées»</i>	152
4.2. L'origine des formes surcomposées.	155
4.3. Surcomposé général vs surcomposé régional	156
4.4. La rareté des formes surcomposées	158
4.5. Auxiliaire vague ou polysémique?	158
4.6. Les formes surcomposées et la voix	161
4.7. Synthèse	164
5. Les phases aspectuelles	164
5.1. Les phases naturelles et leur manifestation en français	165
5.2. L'accompli	166
5.3. Le progressif	169
5.4. Le prospectif	170
5.5. L'expression des bornes temporelles des faits.	172
5.6. Y a-t-il une catégorie flexionnelle de phase aspectuelle?	173
5.7. Synthèse	176
6. Le temps	177
6.1. Passé simple vs imparfait.	178
6.1.1. <i>L'hypothèse de la durée</i>	179
6.1.2. <i>L'hypothèse de la perfectivité</i>	181
6.1.3. <i>L'hypothèse de la progression narrative</i>	185
6.1.4. <i>L'approche reichenbachienne</i>	187
6.2. Présent vs imparfait	193
6.3. Caractérisation sémantique des «formes temporelles» du français.	195
6.3.1. <i>Le présent</i>	196
6.3.2. <i>Le passé simple et le passé composé</i>	197
6.3.3. <i>Le futur simple</i>	198
6.3.4. <i>L'imparfait</i>	199
6.3.5. <i>Le plus-que-parfait</i>	200

6.3.6. <i>Le conditionnel</i>	200
6.4. Le temps, c'est relatif.	202
6.4.1. <i>Temps absolus, relatifs et absolus-relatifs</i>	202
6.4.2. <i>La catégorie flexionnelle de temps</i>	205
6.4.3. <i>La catégorie flexionnelle de décalage</i>	206
6.4.4. <i>Les acceptions secondaires des grammèmes de temps et de décalage</i>	213
6.5. Synthèse: les catégories flexionnelles de temps et de décalage	219
7. Le genre, le nombre et la personne	220
8. Synthèse: les catégories flexionnelles et leurs grammèmes	220

VII. La grammaire d'unification Sens-Texte **223**

1. L'architecture générale de GUST	224
2. La forme des structures de GUST	226
3. Le système des polarités	229
4. L'unification de structures polarisées	233
5. Les modules de GUST et leur articulation	236
5.1. Les grammaires de bonne formation	237
5.1.1. <i>La grammaire sémantique</i>	239
5.1.2. <i>La grammaire syntaxique</i>	242
5.1.3. <i>La grammaire d'arbres</i>	248
5.2. L'articulation des modules.	251
5.3. Les grammaires d'interface	254
5.3.1. <i>La notion de correspondance</i>	255
5.3.2. <i>L'interface sémantique-syntaxe</i>	256
6. La décomposition sémantique en GUST.	267
7. Synthèse.	278

VIII. Un modèle formel de la conjugaison française **281**

1. La grammaire de bonne formation sémantique	283
1.1. Les caractéristiques formelles de $G_{\text{sém}}$	284
1.2. Le typage des actants sémantiques	285
1.3. Les sens temporels	286
1.4. Les sens phasiques / aspectuels	289
1.5. Les sens modaux	295
2. La grammaire de bonne formation syntaxique	298
2.1. Les caractéristiques formelles de G_{synt}	299
2.2. La grammaire d'arbres.	300

2.3. Les règles lexicales	301
2.3.1. <i>La combinatoire des marqueurs de phase aspectuelle</i>	303
2.3.2. <i>La combinatoire des auxiliaires de temps</i>	305
2.3.3. <i>Un auxiliaire du passif?</i>	306
2.4. Les règles grammémiques	307
2.4.1. <i>La combinatoire des grammèmes superficiels de finitude</i>	307
2.4.2. <i>La combinatoire des grammèmes superficiels de mode</i>	308
2.4.3. <i>La combinatoire des grammèmes superficiels de temps</i>	309
2.4.4. <i>La combinatoire des grammèmes superficiels de décalage</i>	310
2.4.5. <i>La combinatoire des grammèmes d'accord</i>	310
2.5. Les règles sagittales	311
2.5.1. <i>Les relations syntaxiques de type auxiliaire-auxilié</i>	311
2.5.2. <i>La relation syntaxique copule-attribut</i>	313
2.5.3. <i>Les relations syntaxiques en dehors du noyau verbal</i>	315
3. La grammaire d'interface sémantique-syntaxe	316
3.1. Les caractéristiques formelles de $I_{\text{sém-synt}}$	318
3.2. Les signes asémantiques	319
3.2.1. <i>Les grammèmes d'accord</i>	319
3.2.2. <i>Les signes régis</i>	321
3.2.3. <i>Les grammies liées à la structure communicative</i>	322
3.2.4. <i>La copule</i>	325
3.3. Les signes de phase aspectuelle et le résultatif	327
3.3.1. <i>Le résultatif</i>	328
3.3.2. <i>Les locutions prépositionnelles de phase aspectuelle</i>	328
3.3.3. <i>Les auxiliaires de phase aspectuelle</i>	331
3.4. Les grammies à valeur temporelle	334
3.4.1. <i>L'acception de base des grammies de temps</i>	334
3.4.2. <i>L'acception de base des grammies de décalage</i>	335
3.5. Les grammies à valeur modale	341
4. Synthèse	347
IX. Conclusion	349
Index des auteurs	355
Index des notions	358
Bibliographie	364
Annexe A. Les formes de PRENDRE	xx
Annexe B. L'encodage XML des grammaires GUST	xxiv

Liste des tableaux

Tableau I — La combinatoire morphologique des auxiliaires	30
Tableau II — Les constructions à deux auxiliaires possibles	31
Tableau III — La combinatoire morphologique des formes à deux auxiliaires	32
Tableau IV — Les constructions à trois auxiliaires possibles	34
Tableau V — La combinatoire morphologique des formes à trois auxiliaires	34
Tableau VI — Le classement traditionnel des formes	74
Tableau VII — Les temps absolus et relatifs dans la tradition	78
Tableau VIII — Le classement des formes selon Imbs	82
Tableau IX — Le classement des formes selon Martinet	87
Tableau X — Le classement des formes selon Touratier	93
Tableau XI — Le classement des formes selon Wilmet	100
Tableau XII — Le classement des formes selon Wilmet revisité	104
Tableau XIII — Le classement des formes selon Gosselin	110
Tableau XIV — Les interprétations de «être V-é»	127
Tableau XV — Les formes surcomposées théoriquement possibles	146
Tableau XVI — La combinatoire des marqueurs de phase	174
Tableau XVII — L'hypothèse des temps à valeur aspectuelle intrinsèque	184
Tableau XVIII — Les relations simples <i>E/R</i> et <i>R/S</i> selon Reichenbach	188
Tableau XIX — Les relations complexes <i>E/R/S</i> selon Reichenbach	188
Tableau XX — Le modèle de Te Winkel	193
Tableau XXI — Les sens du présent et de l'imparfait dans quelques grammaires de référence	193
Tableau XXII — Un système temporel classique à catégories concurrentes	203
Tableau XXIII — Un découpage des signes temporels du français	204
Tableau XXIV — Notre système temporel à deux dimensions	211
Tableau XXV — Le produit des polarités de GUST	231
Tableau XXVI — Tableau récapitulatif des modules de GUST et de leurs objets	279
Tableau XXVII — La combinatoire sémantique des marqueurs de phase aspectuelle	290
Tableau XXVIII — Deux catégories flexionnelles complémentaires	351

Liste des figures

Figure 1 — L'architecture d'un modèle Sens-Texte	37
Figure 2 — La SSém de (7)	39
Figure 3 — La SSyntP de (7)	41
Figure 4 — La SSyntS de (7)	42
Figure 5 — La SMorphP de (7)	43
Figure 6 — La SMorphS de (7)	43
Figure 7 — La SPhonP de (7)	43
Figure 8 — La SPhonS de (7)	43
Figure 9 — Une règle du module syntaxique de surface	44
Figure 10 — Le grammème comme intermédiaire entre sens et formes	48
Figure 11 — Les catégories flexionnelles dans le modèle traditionnel	74
Figure 12 — Les trois axes temporels de Imbs	80
Figure 13 — Les catégories flexionnelles dans le modèle de Imbs	82
Figure 14 — Les catégories flexionnelles dans le modèle de Martinet	86
Figure 15 — Le modèle de Martinet revu par Menanteau	90
Figure 16 — Les catégories flexionnelles dans le modèle de Touratier	93
Figure 17 — Les catégories flexionnelles dans le modèle de Wilmet	99
Figure 18 — Les catégories flexionnelles dans le modèle de Wilmet revisitées	104
Figure 19 — Les catégories flexionnelles dans le modèle de Gosselin	110
Figure 20 — Le découpage du temps réel en trois phases par un fait	165
Figure 21 — La phase accomplie est simultanée au moment d'énonciation	167
Figure 22 — La phase accomplie est postérieure au moment d'énonciation	167
Figure 23 — Caractérisation sémantique du présent	196
Figure 24 — Caractérisation sémantique du passé	197
Figure 25 — Caractérisation sémantique du futur simple	198
Figure 26 — Caractérisation sémantique de l'imparfait	199
Figure 27 — Caractérisation sémantique du plus-que-parfait	200
Figure 28 — Caractérisation sémantique du conditionnel	201
Figure 29 — La combinatoire des grammèmes superficiels verbaux en français	221
Figure 30 — Deux règles simples de GUST	227
Figure 31 — Une structure GUST et sa représentation graphique explicite	229
Figure 32 — Une structure polarisée et sa représentation graphique	230
Figure 33 — Une représentation graphique simplifiée pour une structure polarisée	231
Figure 34 — Deux structures polarisées à combiner	235
Figure 35 — Le résultat de la combinaison de deux structures polarisées	235
Figure 36 — La combinaison partielle de deux structures polarisées	236
Figure 37 — Une grammaire sémantique générant une structure bien formée	240
Figure 38 — Une grammaire sémantique générant une structure mal formée	240
Figure 39 — Une méthode simple pour le typage des actants	241
Figure 40 — Une procédure implicite pour l'unification de types et sous-types	242
Figure 41 — La représentation d'un nœud syntaxique et de l'information grammaticale qui lui est rattachée	244

Figure 42 — Une grammaire syntaxique générant une structure bien formée . . .	246
Figure 43 — Une structure syntaxique mal formée.	248
Figure 44 — Une grammaire d'arbres.	249
Figure 45 — La représentation graphique des structures multipolarisées.	250
Figure 46 — Une grammaire syntaxique à deux polarités	251
Figure 47 — Une grammaire syntaxique articulée	252
Figure 48 — Une grammaire d'arbres articulée	253
Figure 49 — Deux grammaires articulées produisant une structure syntaxique.	253
Figure 50 — La position de la grammaire d'interface sémantique-syntaxe	256
Figure 51 — Une grammaire d'interface sémantique-syntaxe générant une structure de correspondance	258
Figure 52 — La mise en correspondance d'un grammème et de son sens	259
Figure 53 — Une correspondance multivoque	259
Figure 54 — Une grammaire sémantique articulée.	261
Figure 55 — Une interface sémantique-syntaxe articulée.	261
Figure 56 — Une grammaire syntaxique articulée	262
Figure 57 — Une structure de correspondance	262
Figure 58 — Une structure sémantique à neutraliser	263
Figure 59 — Une structure sémantique à mettre en correspondance	263
Figure 60 — Une structure syntaxique à valider.	264
Figure 61 — Deux grammèmes à mettre en correspondance	265
Figure 62 — Une règle d'équivalence sémantique de la TST.	268
Figure 63 — Trois paraphrases qui n'ont pas la même RSém	268
Figure 64 — La décomposition sémantique de 'Torontoise'.	271
Figure 65 — La décomposition sémantique de 'rousse' et de 'femme'.	271
Figure 66 — Les sens « primitifs » dans la grammaire sémantique.	272
Figure 67 — Une représentation sémantique décomposée à réaliser	273
Figure 68 — La réduction de 'personne (de sexe) féminin'	273
Figure 69 — Une règle de réancrage du gouverneur.	274
Figure 70 — La réduction de 'personne (de sexe) féminin' avec réancrage des gouverneurs.	274
Figure 71 — La réduction de 'femme rousse' avec réancrage des gouverneurs	275
Figure 72 — Le typage sémantique des actants en GUST	276
Figure 73 — Un aperçu de la hiérarchie des étiquettes sémantiques du <i>Lexique actif du français</i>	286
Figure 74 — La classification des faits selon Mourelatos.	286
Figure 75 — Les sémantèmes 'avant', 'après' et 'simultané' (R1 à R3)	287
Figure 76 — Les sémantèmes temporels en contexte	288
Figure 77 — Les sémantèmes 'avoir _{acc} ', 'en train' et 'aller' (R4 à R6)	291
Figure 78 — Les sémantèmes 'sur le point' et 'venir' (R7 et R8).	291
Figure 79 — Les sémantèmes de phase aspectuelle en contexte.	292
Figure 80 — La situation dénotée par <i>Ils vont avoir bien dormi</i>	292
Figure 81 — Un sémantème de phase aspectuelle actant de 'penser'	293
Figure 82 — L'interaction sémantique des temps et des phases aspectuelles.	294
Figure 83 — Le réancrage des gouverneurs sémantiques (R9).	294
Figure 84 — Le sémantème 'résultatif' (R10).	295

Figure 85 — Le sémantème ‘manière’ (R11)	296
Figure 86 — Le sémantème ‘devoir’ (R12)	296
Figure 87 — Le sémantème ‘ordonner’ (R13)	297
Figure 88 — Le sémantème ‘supposer’ (R14)	297
Figure 89 — Le sémantème ‘il paraît’ (R15)	298
Figure 90 — Le sémantème ‘en rétrospective’ (R16)	298
Figure 91 — La grammaire d’arbres (R17 et R18)	301
Figure 92 — Un règle demandant une finitude pour un verbe (R19)	302
Figure 93 — Une représentation alternative des parties du discours (RT1 à RT3)	302
Figure 94 — La structure syntaxique de trois locutions prépositionnelles	304
Figure 95 — Quelques règles lexicales triviales (R20 à R22)	304
Figure 96 — La combinatoire des auxiliaires de phase aspectuelle (R23 à R28)	305
Figure 97 — La combinatoire des auxiliaires de temps (R29 et R30)	306
Figure 98 — La règle lexicale pour ÊTRE _{cop} (R31)	306
Figure 99 — La combinatoire des grammèmes de finitude (R32 à R35)	307
Figure 100 — La combinatoire des grammèmes de mode (R36 à R39)	308
Figure 101 — La combinatoire des grammèmes de temps (R40 et R42)	310
Figure 102 — La combinatoire des grammèmes de décalage (R43 et R44)	310
Figure 103 — La combinatoire des grammèmes d’accord (R45 et R49)	311
Figure 104 — La relation syntaxique «Aux» (R50)	312
Figure 105 — La relation syntaxique «Inf» (R51 et R52)	314
Figure 106 — La relation syntaxique «Cop» (R53 et R54)	315
Figure 107 — La relation syntaxique «Prép» (R55 et R56)	316
Figure 108 — La relation syntaxique «Suj» (R57)	317
Figure 109 — L’accord verbe-sujet en nombre (R58 à R60)	320
Figure 110 — L’accord verbe-sujet en personne (R61 à R63)	320
Figure 111 — Le régime de VOULOIR (R64 et R65)	321
Figure 112 — Le régime de PENSER (R66)	322
Figure 113 — Le sommet syntaxique (R67)	322
Figure 114 — Le passif selon Kahane (RT4)	323
Figure 115 — Le participe passé à valeur de passif (R68)	324
Figure 116 — Le participe présent (R69)	324
Figure 117 — La copule selon Kahane (RT5)	325
Figure 118 — La dérivation d’un passif avec copule dans $I_{\text{sém-synt}}$	326
Figure 119 — La copule dans $I_{\text{sém-synt}}$ (R70)	327
Figure 120 — Une double correspondance pour les constructions à copule	327
Figure 121 — Le participe passé à valeur résultative dans $I_{\text{sém-synt}}$ (R71)	329
Figure 122 — Les prépositions ‘À CÔTÉ’ et SUR dans $I_{\text{sém-synt}}$ (R72 et R73)	329
Figure 123 — Les prépositions ‘EN TRAIN’ et ‘SUR LE POINT’ dans $I_{\text{sém-synt}}$ (R74 et R75)	330
Figure 124 — La phase accomplie dans $I_{\text{sém-synt}}$ (R76 et R77)	331
Figure 125 — La phase prospective et le passé récent dans $I_{\text{sém-synt}}$ (R78 et R79)	332
Figure 126 — Les étapes de la dérivation de <i>Ils vont être en train de dormir</i>	333

Figure 127 — L'expression des sens 'simultané' et 'postérieur' (R80 et R81) . . .	335
Figure 128 — L'expression du sens 'antérieur' dans $I_{\text{sém-synt}}$ (R82 à R84) . . .	336
Figure 129 — L'acception de base des grammèmes de décalage (R85 à R87) .	336
Figure 130 — La représentation partielle de (175) dans $I_{\text{sém-synt}}$	338
Figure 131 — La représentation complète de (175) dans $I_{\text{sém-synt}}$	341
Figure 132 — Le gérondif dans $I_{\text{sém-synt}}$ (R88)	342
Figure 133 — L'infinitif de consigne $I_{\text{sém-synt}}$ (R89)	343
Figure 134 — L'impératif dans $I_{\text{sém-synt}}$ (R90)	343
Figure 135 — Le futur de supposition dans $I_{\text{sém-synt}}$ (R91)	344
Figure 136 — Le conditionnel de réserve dans $I_{\text{sém-synt}}$ (R92)	344
Figure 137 — Les collocations lexicales dans $I_{\text{sém-synt}}$ (RT6 et RT7)	345
Figure 138 — Le futur antérieur rétrospectif dans $I_{\text{sém-synt}}$ (R93)	346

Conventions typographiques

[...]	délimitation d'un exemple inséré dans le texte
(québ.)	langue ou registre auquel appartient un exemple
fonction structurante	terme introduit ou défini pour la première fois
infinitif	grammème
–ER	morphème suffixal
TRINQUER ₁	lexie (avec indice distinctif)
「TIRER LE DIABLE PAR LA QUEUE」	locution
TRINQUER	vocable
‘triquer ₁ ’	sémantème (avec indice distinctif)
<i>triquer</i>	signifiant écrit d'un signe linguistique
\mathcal{P}	ensemble
$G_{\text{sém}}$	module d'une grammaire GUST
$\text{noeud}_{\text{sém}}$	type d'objet d'une structure GUST
<i>étiquette</i>	fonction d'une structure GUST

Symboles utilisés

⊕	composition de deux signes linguistiques
⊔	unification de deux structures
≡	équivalence entre deux structures
↔	correspondance entre deux structures
	disjonction dans les structures

Liste des abréviations

acc	accompli
act.	actif
ant	antérieur
arg	argot français
aux	auxiliaire
c. nom	complément du nom
cond	conditionnel
cop	copule
déc	décalé
dép	dépendant
dét	déterminant
FA	futur antérieur
fém.	féminin
fin.	finitude
FUT	futur
futur s	futur simple
G_{arbre}	grammaire d'arbres
GÉR	gérondif
G_{phon}	grammaire de bonne formation phonologique
$G_{\text{sém}}$	grammaire de bonne formation sémantique
G_{synt}	grammaire de bonne formation syntaxique
G_{topo}	grammaire de bonne formation morphotopologique
GUP	grammaire d'unification polarisée
GUST	grammaire d'unification Sens-Texte
HPSG	grammaire syntagmatique dirigée par les têtes
imp	imparfait
impér	impératif
INACC	inaccompli
ind	indicatif
inf.	infinitif
$I_{\text{sém-synt}}$	grammaire d'interface sémantique-syntaxe
$I_{\text{synt-topo}}$	grammaire d'interface syntaxe-morphotopologie
$I_{\text{topo-phon}}$	grammaire d'interface morphotopologie-phonologie
LFG	grammaire lexicale fonctionnelle
masc	masculin
N	nom
nbre	nombre
n-déc	non-décalé
PA	passé antérieur
part.	participe
part-ant.	participe-ant (participe présent)
part-é	participe-é (participe passé)
passé s	passé simple

PC	passé composé
pdd	partie du discours
pers.	personne
pl	pluriel
pos	position
post.	postérieur
PQP	plus-que-parfait
prép	préposition
prés.	présent
PS	passé simple
quant	quantifieur
québ	français québécois
rég.	français d'une variété régionale
RMorphP	représentation morphologique profonde
RMorphS	représentation morphologique de surface
RPhonP	représentation phonologique profonde
RPhonS	représentation phonologique de surface
RSém	représentation sémantique
RSyntP	représentation syntaxique profonde
RSyntS	représentation syntaxique de surface
SAnaphP	structure anaphorique profonde
SAnaphS	structure anaphorique de surface
sg	singulier
sim	simultané
SMorphP	structure morphologique profonde
SMorphS	structure morphologique de surface
SPhonP	structure phonologique profonde
SPhonS	structure phonologique de surface
SRéf.	structure référentielle
SRhét	structure rhétorique
SSém	structure sémantique
SSém-Comm	structure sémantico-communicative
SSynt-CommP	structure syntaxico-communicative profonde
SSynt-CommS	structure syntaxico-communicative de surface
SSyntP	structure syntaxique profonde
SSynt-ProsP	structure syntaxico-prosodique profonde
SSynt-ProsS	structure syntaxico-prosodique de surface
SSyntS	structure syntaxique de surface
subj.	cas subjectif
subj.	subjonctif
suj.	sujet
T_0	moment d'énonciation
T_1	moment de référence
TAG	grammaire d'arbres adjoints
TST	théorie Sens-Texte
V	verbe

Remerciements

Je tiens tout d'abord à remercier chaleureusement mes sympathiques directeurs, Sylvain Kahane et Igor Mel'čuk. Ce sont de vrais passionnés qui donnent sans compter, mais aussi des tortionnaires impitoyables prêts à tout pour extraire le jus de leur victime (qu'ils iraient jusqu'à pourchasser sur les plages de la Méditerranée). S'il y a du bon dans cette thèse, ils y sont pour quelque chose.

C'est à l'Université de Stuttgart avec Leo Wanner, en 2001, que j'ai naïvement commencé à élaborer un modèle de la flexion verbale en français. De retour à Montréal, j'ai abondamment discuté du sujet avec Joel Altman et Michel Usereau, ce qui m'a poussé à faire du temps verbal le sujet de ma thèse. À Paris, au début de ce doctorat, j'aurais dormi sous les ponts si je n'avais pas été accueilli si gentiment chez Hiyon Yoo et Kim Gerdes, qui ne m'a pas laissé tranquille avant d'avoir bien compris ce que je cherchais à faire dans mon travail. Le convaincre m'a aidé à me faire les dents. Plus tard, mon passage au séminaire Talana de l'Université Paris 7 a été l'occasion de revoir mon modèle et d'aboutir à la version plus ou moins finale. J'ai également bénéficié des commentaires de Lucie Barque, Dina El Kassas, Kim Gerdes, Jasmina Milićević et Céline Raynal, qui ont bien voulu relire certains passages de cette thèse. Lors de l'examen de synthèse, les questions et remarques des professeurs Yves-Charles Morin et Alain Polguère m'ont été très utiles. Enfin, les commentaires du jury lors de la soutenance m'ont permis d'améliorer le texte.

Cette thèse a été rendue possible grâce au soutien financier du Conseil de la recherche en sciences humaines du Canada (bourse de doctorat # 752-02-1501), de l'Université de Montréal (Département de linguistique, Fondation J.A. de Sève), du Gouvernement français (soutien aux cotutelles), du Ministère des relations internationales du Québec (soutien aux cotutelles), et du bureau des relations internationales de l'Université Paris 7 (soutien aux cotutelles).

Enfin, sans le soutien continu de ma famille et de mes amis, qui ont enduré mes angoisses et mes absences, je n'aurais jamais trouvé l'énergie de terminer. Merci!

I. Introduction

Les humains, en plus de percevoir l'espace qui les entoure, sont conscients du temps qui s'écoule. La dimension temporelle de l'univers dans lequel nous vivons est directement liée à divers phénomènes d'une importance cruciale pour notre survie même : les liens de cause à effet, la mémoire et l'apprentissage, notre propre vieillissement, etc. Il semble donc tout à fait naturel que le temps physique trouve son écho dans les langues que nous parlons. Il est même difficile d'imaginer une langue humaine qui n'offrirait aucun moyen de situer les faits¹ dans le temps lorsque cela est nécessaire. En fait, beaucoup de langues vont même jusqu'à forcer les locuteurs à le faire. C'est le cas notamment du français, où les verbes à l'indicatif, qui constituent le cœur de la plupart des énoncés, doivent porter des marques de temps grammatical qui sont le reflet linguistique du temps physique. C'est ce phénomène que nous nous proposons d'étudier dans cette thèse.

1 Problématique de la recherche

De façon générale, nous nous intéressons à la modélisation formelle de la langue en synchronie, en particulier celle de la grammaire (par opposition à celle du lexique). Plus précisément, ce sont les problèmes du temps grammatical² et de l'aspect en français qui retiennent notre attention, c'est-à-dire les signes flexionnels qui servent essentiellement à situer les faits dans le temps physique ou à spécifier la nature du lien entre ces faits et leurs bornes temporelles intrinsèques. Ainsi, nous nous proposons de produire un modèle fonctionnel formel des signes flexionnels de la conjugaison française contemporaine. Ce modèle doit bien s'intégrer dans une description globale de la langue et doit mettre en relation les sens grammaticaux de nature temporelle et les formes qui les expriment (ou plus précisément, la représentation de ces sens et de ces formes).

-
1. Dans cette thèse, nous utilisons le terme *fait* plutôt que *procès* pour parler du référent des verbes.
 2. À moins d'indication contraire, afin d'alléger le texte, nous utiliserons dorénavant le terme *temps* dans le sens de *temps grammatical*.

Notre recherche s'articule autour de quatre axes principaux : théorie, méthodologie, description et formalisation. Nous aurions aimé aborder également les problèmes liés à l'implémentation informatique de notre modèle linguistique, mais nous avons dû y renoncer par manque de temps.

Au niveau théorique, nous cherchons à mieux comprendre la nature des signes grammaticaux en général, et comment ils s'organisent en un système flexionnel. Sur ce point, nous verrons que la littérature existante fournit déjà des réponses globalement satisfaisantes.

Au niveau méthodologique, nous voulons définir des critères pour guider la construction d'un modèle fonctionnel de la flexion (en particulier verbale). Notamment, nous cherchons à développer une approche permettant d'éviter les pièges de la polysémie et de la phraséologie dans le domaine flexionnel. Nous verrons que la lexicographie offre des solutions intéressantes pour notre étude.

Au niveau descriptif, les questions qui sous-tendent notre recherche sont les suivantes :

- 1) Quelles sont les catégories flexionnelles pertinentes pour le verbe en français ?
- 2) Quels en sont les grammèmes, et quels sens expriment-ils ?
- 3) En particulier, comment les grammèmes verbaux du français permettent-ils de situer les faits dans le temps physique ?

L'hypothèse principale que nous tentons de démontrer est qu'il n'y a pas qu'une seule catégorie flexionnelle de temps en français (qui comprendrait le présent, l'imparfait, le futur simple, etc.), mais bien deux : une, que nous appelons «décalage», qui situe un point de référence par rapport au moment de l'énonciation et une autre, que nous appelons «temps», qui situe les faits dénotés par les verbes par rapport à ce point de référence. Par exemple, le conditionnel dans *Il a dit qu'il viendrait* se décompose en un grammème de temps **postérieur** (exprimé par –R–) et un grammème de décalage **décalé** (exprimé par –AI–).

Enfin, au niveau formel, nous poursuivons le développement du formalisme de la grammaire d'unification Sens-Texte. Nous proposons des outils formels pour la description des langues, notamment au niveau sémantique. Nous tenterons d'encoder notre description de la conjugaison française dans ce formalisme afin d'en tester les limites. Le modèle que nous développons est un modèle fonctionnel de la langue pouvant fonctionner autant comme un modèle génératif que transductif (en synthèse ou en analyse).

2 Pourquoi le temps grammatical ?

Le problème du temps grammatical, ou même de la flexion verbale dans son ensemble, ne représente qu'une infime partie de la grammaire du français. Cependant, ce n'est aucunement un problème marginal, puisque le verbe est omniprésent dans cette langue. Le discours normal en français est truffé de verbes fléchis en temps. En ce sens, les signes grammaticaux ayant trait à la situation temporelle des faits, bien qu'ils ne forment qu'une toute petite partie des signes linguistiques, se trouvent au cœur de la langue (et c'est d'ailleurs le cas dans beaucoup d'autres langues, bien que les systèmes grammaticaux varient considérablement). Le problème de la flexion verbale en général, et notamment celui du temps grammatical, doit donc faire partie des principaux phénomènes à décrire.

Par ailleurs, à travers l'étude de la flexion verbale en français, c'est le développement du formalisme de la grammaire d'unification Sens-Texte qui nous intéresse. Nous croyons que ce phénomène se prête bien au jeu, puisqu'il implique tous les niveaux de fonctionnement de la langue, de la sémantique à la phonologie. Toutefois, nous laisserons de côté les problèmes morphologiques et phonologiques liés à la réalisation de surface des morphèmes verbaux pour nous concentrer plutôt sur la conjugaison dans l'interface sémantique-syntaxe.

3 Organisation de la thèse

Le Chapitre II délimite les phénomènes à l'étude dans le cadre de ce travail. Nous dressons méthodiquement une liste de toutes les formes verbales à considérer et nous écartons celles qui d'emblée ne sont pas pertinentes pour notre recherche.

Au Chapitre III nous présentons brièvement la théorie linguistique qui soutient notre travail, la théorie Sens-Texte.

On trouvera au Chapitre IV une discussion des notions de grammème, grammie et catégorie flexionnelle. Nous y proposons une méthodologie inspirée de la lexicographie pour l'étude des signes grammaticaux.

Au Chapitre V nous passons en revue quelques modèles de la conjugaison qui ont été proposés par nos prédécesseurs. Nous tenterons de traduire ces modèles en termes de grammèmes au sens où nous l'entendons au Chapitre IV, et nous verrons que cela n'est pas possible.

C'est au Chapitre VI, qui constitue le cœur descriptif de la thèse, que nous analysons le système flexionnel verbal du français. Nous y identifions les grammèmes verbaux du français ainsi que leurs principales acceptions.

Le Chapitre VII prépare le terrain à la formalisation du modèle élaboré au chapitre précédent. Nous y présentons et développons la grammaire d'unification Sens-Texte, qui nous sert de cadre formel.

On trouvera au Chapitre VIII le modèle formel que nous proposons pour la conjugaison française.

Enfin, au Chapitre IX, nous conclurons par un bref résumé des principaux éléments de la thèse et proposerons quelques pistes à explorer pour résoudre les problèmes laissés en suspens.

À la fin de chaque chapitre, ainsi que des sections les plus importantes, nous faisons une brève synthèse des principaux éléments présentés.

II. Délimitation du champ d'étude

Avant toute chose, il convient de délimiter les faits linguistiques que nous cherchons à décrire. Précisons d'emblée que nous n'avons pas la prétention de décrire l'ensemble des phénomènes liés à la conjugaison des verbes. Nous ne nous intéressons, dans le cadre de cette thèse, qu'aux faits liés à l'expression des relations temporelles dans la conjugaison en français standard contemporain, et ce, seulement du point de vue de l'interface sémantique-syntaxe. Par conséquent, nous *ne* traitons *pas*, entre autres, les phénomènes suivants :

- les marqueurs temporels non verbaux (*cf.* Vet 1980 ou Gosselin 1996);
- la voix grammaticale (voir à ce sujet Mel'čuk 2001b);
- la syntaxe des auxiliaires (voir à ce sujet Abeillé & Godard 2002b);
- les considérations strictement morphologiques, morpho-phonologiques ou phonologiques (voir à ce propos Van Den Eynde & Blanche-Benveniste 1970, Morin 1987, Touratier 1996, ou encore Le Goffic 1997);
- la variation historique (*cf.* par exemple Cornu 1953 ou Morin 2001) ou dialectale (voir par exemple Jolivet 1984 ou Laurendeau 2000) de la flexion verbale.

Nous serons cependant appelé à discuter dans une certaine mesure les phénomènes suivants, soit parce qu'ils sont sémantiquement liés au temps grammatical, soit parce qu'ils interagissent directement avec les signes temporels :

- les marqueurs «aspectuels», en raison de leur sens intimement lié au temps grammatical;
- le mode et la finitude, puisqu'ils peuvent bloquer ou encore forcer l'expression du temps grammatical;
- de façon moins élaborée, les questions d'Aktionsart, seulement là où ce phénomène a une incidence sur l'expression du temps grammatical.

Nous proposons dans les pages qui suivent un inventaire systématique des formes verbales (synthétiques et analytiques) que nous allons considérer dans cette thèse. Les formes les plus simples sont déjà répertoriées dans les ouvrages plus ou

moins traditionnels sur la conjugaison française. Nous utilisons notamment le *Bescherelle* (1998), le *Larousse de la conjugaison* (1999) ainsi que le *Les formes conjuguées du verbe français oral et écrit* de Le Goffic (1997). Ces ouvrages fournissent toutes les formes synthétiques existantes et les formes analytiques de base. À partir de l'ensemble de ces données nous construisons de façon systématique les formes analytiques en combinant les auxiliaires (par exemple, *allait avoir été enseveli*) pour obtenir l'inventaire complet des formes verbales.

Il n'est pas toujours évident de distinguer les simples combinaisons de lexies des phénomènes qui relèvent de la flexion. Nous devons donc anticiper quelque peu sur les chapitres qui suivent pour justifier certains de nos choix. Toutes les caractérisations sémantiques présentées dans cette section sont volontairement vagues. Il s'agit seulement de premières évaluations grossières et nous reviendrons sur ces différentes constructions dans les chapitres suivants. En cas de doute, nous inclurons les constructions suspectes dans notre liste, quitte à les rejeter plus tard une fois que nous aurons présenté les critères qui dictent nos choix.

1 Précision terminologique

Nous utilisons les termes de la grammaire traditionnelle pour faire référence aux différentes formes verbales («présent», «imparfait», «passé composé», etc.). Ces formes ne correspondent jamais à un seul grammème. Par exemple, nous ne pouvons pas le justifier ici puisque nous n'avons pas encore présenté les concepts nécessaires, mais nous considérons que ce que la grammaire traditionnelle appelle «présent» exprime deux grammèmes temporels, **simultané** et **décalé**, en plus des grammèmes de mode **indicatif** et de finitude **fini**. D'autre part, les termes utilisés traditionnellement ne correspondent pas forcément à des grammèmes de notre modèle (nous n'avons pas de grammème ***imparfait** ou ***conditionnel**, par exemple). Par convention, nous utiliserons toujours une police spéciale lorsque nous parlerons d'un grammème, et une police normale lorsque nous ferons référence à une forme verbale selon la terminologie traditionnelle¹.

2 Les formes synthétiques

Les formes synthétiques (aussi appelées «formes simples») sont bien connues et appellent peu de commentaires :

- Le présent: *savez*.
- L'imparfait: *saviez*.
- Le futur simple: *saurez*.
- Le passé simple: *sûtes*.
- Le conditionnel: *sauriez*.
- Le subjonctif: *sachiez*.
- L'impératif: *sachez*.
- L'infinitif: *savoir*.
- Le participe présent: *sachant*.
- Le participe passé: *su*².

On remarquera qu'il manque à cette liste l'imparfait du subjonctif [*sussiez*]. En effet, s'il est encore à l'occasion utilisé à l'écrit par certains auteurs (et même à peu près compris par certains de leurs lecteurs), il n'est en général pas maîtrisé activement par les locuteurs natifs, même dans les milieux les plus scolarisés³. Il appartient donc à un autre code que celui du français contemporain et nous l'excluons par conséquent de notre domaine d'étude. Le fait qu'on puisse encore le trouver dans quelques textes contemporains ne constitue pas un contre-argument suffisant. Il est également toujours possible d'utiliser des tournures archaïques comme *Je ne me puis dépêtrer de cet homme* (Lafontaine, *Le Cocu, battu, et content*), mais on ne dit pas pour autant que ces tournures sont encore vivantes aujourd'hui.

-
1. Nous recommandons fortement au lecteur de consulter la liste des conventions typographiques au début de cette thèse (p. xvii).
 2. Les verbes ne peuvent pas tous être utilisés au participe passé sans faire partie d'une forme analytique. Ce n'est possible que si le verbe est compatible avec la voix passive ou s'il peut avoir une valeur de résultatif (cf. Ch. VI, § 1.2.2, p. 124).
 3. Nous ne disposons pas de données objectives pour appuyer cette affirmation, mais le fait nous semble suffisamment évident. Une petite enquête informelle dans notre entourage nous conforte dans cette idée.

II. Délimitation du champ d'étude

3 Les formes analytiques à un auxiliaire

Les formes analytiques (aussi appelées « formes composées » quand ce sont des verbes) sont constituées de plus d'un mot-forme. Les formes analytiques verbales sont composées d'un verbe précédé d'un auxiliaire ou semi-auxiliaire (la distinction entre les deux ne nous intéresse pas ici). La plupart des grammaires modernes reconnaissent à peu près les mêmes constructions à valeur temporelle ou aspectuelle, soit :

- Le passé composé : « AVOIR + V_{pp} » ou « ÊTRE + V_{pp} » selon la lexie en jeu ou le contexte syntaxique (par exemple, la voix pronominale impose ÊTRE).
- Le futur périphrastique : « ALLER + V_{inf} ».
- Le passé récent : « VENIR + V_{inf} ».

Deux autres constructions sont fréquemment mentionnées dans les ouvrages de référence, mais il n'est pas toujours tout à fait clair si elles sont considérées comme faisant partie du paradigme de la conjugaison ou non :

- Un marqueur d'aspect progressif : « ⁷EN TRAIN⁷ + V_{inf} ».
- Un marqueur d'imminence : « ⁷SUR LE POINT⁷ + V_{inf} »

Il semble que la copule ÊTRE, contrairement à ce qui est affirmé dans toutes les grammaires consultées, ne fasse pas partie des constructions « ⁷EN TRAIN⁷ + V_{inf} » et « ⁷SUR LE POINT⁷ + V_{inf} », puisque celles-ci peuvent s'utiliser sans cette copule :

- (1) *Je l'ai surprise en train de fouiller dans mes affaires.*
- (2) *Sur le point d'éclater en sanglots, il alla se réfugier dans les toilettes.*

Ces constructions ont une distribution semblable aux prépositions :

- (3) a. *Mes clés étaient sous la table.*
b. *Ma sœur était en train de fouiller dans mes affaires <était sur le point de partir>.*
- (4) a. *J'ai trouvé mes clés sous la table.*
b. *J'ai trouvé ma sœur en train de fouiller dans mes affaires <sur le point de partir>.*

D'autre part, elles peuvent apparaître sans dépendant verbal, ce qui contraste fortement avec les auxiliaires authentiques⁴:

- (5) a. — *Est-ce que tu as écrit à Pierre?*
 — *J'étais en train <J'étais sur le point>.*
- b. — *Est-ce que tu as écrit à Pierre?*
 — * *Je vais <* J'ai>.*

Enfin, 「EN TRAIN」 peut être remplacé, quand le contexte syntaxique l'exige, par 「EN PLEIN」, qui a la même distribution si ce n'est qu'il apparaît avec des noms plutôt que des verbes⁵:

- (6) a. *Il était en train de réfléchir.*
 b. *Il était en pleine réflexion.*

「EN TRAIN」 et 「SUR LE POINT」 ne sont donc pas des auxiliaires (ni même des verbes) et ne peuvent pas entrer dans la composition de formes analytiques verbales. Néanmoins, vu l'importance apparente de l'expression «「EN TRAIN」 + V_{inf}» dans le système verbal ainsi que son traitement privilégié dans les grammaires en général, nous allons la considérer dans cette thèse. Nous accorderons également un sursis à 「SUR LE POINT」. Pour les fins du présent chapitre, nous allons traiter ces deux locutions comme des auxiliaires afin d'alléger le texte, bien qu'elles n'en soient pas.

Outre les constructions que nous venons de mentionner, (Riegel, Pellat & Rioul 1994) traitent les expressions *commencer à*, *se mettre à*, *finir de*, *achever de* et *cesser de* comme des auxiliaires au même titre que, par exemple, ALLER. Nous croyons plutôt qu'il s'agit de verbes ordinaires utilisés de façon tout à fait libre et compositionnelle. En effet, ils ne se combinent pas seulement aux verbes, mais également aux noms [*commencer à travailler* ~ *commencer le travail*, *finir de travailler* ~ *finir le travail*, etc.], ce qui n'est pas le cas des véritables marqueurs d'«aspect», qui ne peuvent se combiner qu'aux verbes seulement [*il va réussir* ~

4. Le cas de VENIR est à cheval entre les deux: *Je viens juste* est grammatical mais pas * *Je viens*.
5. 「SUR LE POINT」 n'a pas d'équivalent clairement identifiable qui s'utiliserait avec les noms, mais il existe de nombreuses collocations qui expriment le même sens: *sur le point de partir* ~ *sur son départ*, *sur le point de pleurer* ~ *au bord des larmes*, etc.

II. Délimitation du champ d'étude

* *il va la réussite, il vient de finir* ~ * *il vient de la fin*, etc.]. D'autre part, ils se laissent aisément paraphraser par des lexies d'autres parties du discours [*commencer à manger* ~ *le début du repas, finir de travailler* ~ *la fin du travail*, etc.].

Les auxiliaires, puisqu'ils sont des verbes, ainsi que la copule ÊTRE qui accompagne les locutions prépositionnelles «EN TRAIN» et «SUR LE POINT», doivent être fléchis et prendre une des formes synthétiques mentionnées plus haut (§ 2, p. 27). Le tableau suivant résume la combinatoire de ces expressions avec les formes synthétiques (un crochet indique la compatibilité, un astérisque l'incompatibilité).

	prés	imp	futur s.	cond	passé s.	subj	impér	inf	part prés	part passé
AVOIR	✓	✓	✓	✓	✓	✓	✓	✓	✓	*i
ÊTRE	✓	✓	✓	✓	✓	✓	✓ ⁱⁱ	✓	✓	*i
ALLER	✓	✓	*	*	*	*	*	*	*	*
VENIR	✓	✓	✓	✓	✓	✓	*	✓	✓	*
ÊTRE + «EN TRAIN»	✓	✓	✓	✓	✓	✓	*	✓	✓	*iii
ÊTRE + «SUR LE POINT»	✓	✓	✓	✓	✓	✓	*	✓	✓	*iii

Tableau I — La combinatoire morphologique des auxiliaires

- i. N'est possible que dans les formes surcomposées (§ 4, p. 31).
- ii. Sauf pour les verbes pronominaux : *sois parti* est correct mais pas **sois-toi endormi*.
- iii. N'est possible que si la copule porte elle-même un auxiliaire, puisque ÊTRE n'accepte pas le passif (voir note 2, p. 27).

À partir de ces informations, on peut générer la liste complète des formes analytiques à un auxiliaire pour un verbe donné (voir Annexe A, § 2, p. xx).

Les auxiliaires peuvent également se combiner entre eux, ce qui permet de produire des constructions à deux, trois ou quatre auxiliaires.

4 Les formes analytiques à deux auxiliaires

Le tableau suivant présente les paires d'auxiliaires théoriquement possibles ainsi que la grammaticalité de chacune des combinaisons. L'ordre relatif des auxiliaires est pertinent [*il va s'être endormi* ~ **il s'est allé endormir*]. La colonne de titre du tableau donne le premier terme de la paire, la ligne de titre en donne le second terme (par exemple, la première ligne doit se lire comme suit: *avoir eu* (*endormi*), *avoir été* (*parti*), *avoir allé* (*endormir*), *avoir venu* (*d'endormir*), ...). Attention: l'auxiliaire ÊTRE que nous considérons ici est celui qui alterne avec AVOIR dans la formation du passé composé. Il ne faut pas le confondre avec celui qui apparaît dans la voix passive. De façon similaire, l'auxiliaire ALLER ne doit pas être confondu avec le verbe de mouvement homonyme.

	AVOIR	ÊTRE	ALLER	VENIR	「EN TRAIN」	「SUR LE POINT」
AVOIR	✓	✓	*	*	✓	✓
ÊTRE	*	*	*	*	*	*
ALLER	✓	✓	*	✓	✓	✓
VENIR	*	*	*	*	*	*
「EN TRAIN」	*	*	*	*	*	*
「SUR LE POINT」	? ⁱ	? ⁱ	*	*	*	*

Tableau II — Les constructions à deux auxiliaires possibles

- i. ? *Il est sur le point d'avoir fini* <*d'être parti*> nous semble plutôt lourd, mais la construction n'est pas impossible, comme en témoigne cette phrase trouvée sur internet: *Une petite tradition du lieu : lorsque vous êtes sur le point d'avoir fini votre verre, le patron vous en remet un direct.* Dans ces constructions, l'auxiliaire AVOIR / ÊTRE peut être omis [*lorsque vous êtes sur le point de finir...*].

Le premier auxiliaire d'une forme à deux auxiliaires est celui qui porte la flexion. Le tableau suivant présente, pour chaque construction à deux auxiliaires possible, sa combinatoire morphologique.

	prés	imp	futur s.	cond	passé s.	subj	impér	inf	part prés	part passé
AVOIR + AVOIR ⁱ	✓	✓	*	✓	*	*	*	*	*	*
AVOIR + ÊTRE ⁱ	✓	✓	*	✓	*	*	*	*	*	*
AVOIR + 「EN TRAIN」	*	*	*	✓	*	*	*	*	*	*
AVOIR + 「SUR LE POINT」	*	*	*	✓	*	*	*	*	*	*
ALLER + AVOIR	✓	✓	*	*	*	*	*	*	*	*
ALLER + ÊTRE	✓	✓	*	*	*	*	*	*	*	*
ALLER + VENIR	✓	✓	*	*	*	*	*	*	*	*
ALLER + 「EN TRAIN」	✓	✓	*	*	*	*	*	*	*	*
ALLER + 「SUR LE POINT」	✓	✓	*	*	*	*	*	*	*	*
「SUR LE POINT」 + AVOIR ⁱⁱ	✓	✓	✓	✓	✓	✓	*	✓	✓	*iii
「SUR LE POINT」 + ÊTRE ⁱⁱ	✓	✓	✓	✓	✓	✓	*	✓	✓	*iii

Tableau III — La combinatoire morphologique des formes à deux auxiliaires

- i. Nous justifions au Chapitre VI nos jugements sur l'acceptabilité des formes surcomposées (Ch. VI, § 4.1, p. 145).
 ii. Voir note i, Tableau II, p. 31.
 iii. Voir note 2, p. 27.

On trouvera la liste complète des formes à deux auxiliaires à l'Annexe A (§ 3, p. xxii). Il est intéressant de noter que l'auxiliaire AVOIR est le seul à ne pas avoir la même combinatoire morphologique selon qu'il est ou non le seul auxiliaire dans la construction.

Pour les constructions AVOIR + (ÊTRE) 「EN TRAIN」 et AVOIR + (ÊTRE) 「SUR LE POINT」, il y a un fort contraste d'acceptabilité entre les formes au conditionnel et les autres, que nous rejetons. Pour ce qui est des formes surcomposées, il faut faire attention de ne pas confondre les dialectes. Dans certains dialectes, les formes sur-

II. Délimitation du champ d'étude

composées sont beaucoup plus utilisées et elles peuvent avoir une combinatoire morphologique différente⁶. Nous ne considérons ici que le français standard. Dans cette variété de français, il ne faut retenir que les formes surcomposées au présent, à l'imparfait ou au conditionnel. Riegel, Pellat & Rioul (1994) sont d'avis que les deux premières sont meilleures que la dernière :

Les autres formes surcomposées, qui sont parallèles aux formes composées, à l'exception des formes correspondant au passé antérieur de l'indicatif et au plus-que-parfait du subjonctif, sont très rares : il avait eu chanté, il aura eu chanté, il aurait eu chanté, qu'il ait eu chanté, avoir eu chanté, ayant eu chanté. Elles ne sont pas toujours reconnues par un locuteur français standard. Pour cette raison, elles ne figurent pas ici dans les tableaux de conjugaisons.

Riegel, Pellat & Rioul (1994 : p. 152).

Néanmoins, les études de Carruthers (1996) révèlent que le conditionnel surcomposé est aussi attesté.

5 Les formes analytiques à trois auxiliaires

L'auxiliaire qui est au sommet d'une forme à deux auxiliaires peut lui même dépendre d'un autre auxiliaire. Le Tableau IV ci-dessous montre les combinaisons possibles. Il se lit comme suit : *avoir eu eu (dormi)*, *avoir eu été (parti)*, *avoir eu été sur le point (de dormir)*, *avoir allé avoir (dormi)*, etc.

Comme pour les formes à deux auxiliaires, c'est l'auxiliaire qui est le sommet syntaxique de l'expression qui porte la flexion. Le Tableau V ci-dessous présente, pour chaque construction à trois auxiliaires possible, sa combinatoire morphologique.

Ces formes sont données à l'Annexe A (§ 4, p. xxiii). Certaines de ces formes sont très rares et ne sont pas attestées dans les corpus. Elles sont souvent ignorées dans les grammaires.

6. Voir notamment les travaux de Carruthers (1994, 1996, 1999) à ce sujet.

	AVOIR + AVOIR	AVOIR + ÊTRE	AVOIR + 「EN TRAIN」	AVOIR + 「SUR LE POINT」	ALLER + AVOIR	ALLER + ÊTRE	ALLER + VENIR	ALLER + 「EN TRAIN」	ALLER + 「SUR LE POINT」	「SUR LE POINT」 + AVOIR	「SUR LE POINT」 + ÊTRE
AVOIR	*	*i	*	*	*	*	*	*	*	?ii	?ii
ÊTRE	*	*	*	*	*	*	*	*	*	*	*
ALLER	*	*	*	*	*	*	*	*	*	?ii	?ii
VENIR	*	*	*	*	*	*	*	*	*	*	*
「EN TRAIN」	*	*	*	*	*	*	*	*	*	*	*
「SUR LE POINT」	*	*	*	*	*	*	*	*	*	*	*

Tableau IV — Les constructions à trois auxiliaires possibles

- i. Cette construction est possible dans certains dialectes, mais pas en français standard. Voir Carruthers (1994, 1996).
- ii. Voir note i, Tableau II, p. 31.

	prés	imp	futur s.	cond	passé s.	subj	impér	inf	part prés	part passé
AVOIR + 「SUR LE POINT」 + AVOIR ⁱ	*	*	*	?	*	*	*	*	*	*
AVOIR + 「SUR LE POINT」 + ÊTRE ⁱ	*	*	*	?	*	*	*	*	*	*
ALLER + 「SUR LE POINT」 + AVOIR ⁱ	✓	✓	*	*	*	*	*	*	*	*
ALLER + 「SUR LE POINT」 + ÊTRE ⁱ	✓	✓	*	*	*	*	*	*	*	*

Tableau V — La combinatoire morphologique des formes à trois auxiliaires

- i. Voir note i, Tableau II, p. 31.

Nous ne trouvons pas de constructions à quatre auxiliaires dans les corpus et les seules que nous pouvons imaginer sont celles formées de ALLER + AVOIR + 「SUR LE POINT」 + AVOIR/ ÊTRE [^{??}*Il va <allait> avoir été sur point d'avoir fini <d'être parti>*]. Ces constructions sont trop lourdes et nous les excluons de notre étude.

II. Délimitation du champ d'étude

III. La théorie Sens-Texte

Notre travail se situe dans le cadre de la théorie Sens-Texte (TST). En particulier, nous nous appuyons fortement sur l'appareil conceptuel élaboré dans le *Cours de morphologie général* de Mel'čuk (1993, 1994, 1996a, 1997b, 2000). Toutefois, le formalisme que nous utilisons est celui de la grammaire d'unification Sens-Texte. Ce dernier sera exposé en détail au Chapitre VII. Dans le présent chapitre, nous présentons en quelques pages les grandes lignes de la TST.

1 Présentation générale

La TST a vu le jour à l'Académie des sciences de Moscou vers la fin des années 1960 comme une solution au problème de la traduction automatique anglais ↔ russe. Ses textes fondateurs sont ceux de Žolkovskij & Mel'čuk (1965, 1967, 1970) et de Mel'čuk (1973, 1974). On en trouve également de nombreuses introductions, notamment celles de Mel'čuk (1997a), Polguère (1998a), Kahane (2001, 2003a) ou encore Milićević (2006)¹.

Cette théorie est basée sur les trois postulats suivants :

- 1) La langue est conçue comme la mise en correspondance de sens et de formes. Cette correspondance est multivoque dans les deux directions (à un même sens peuvent correspondre plusieurs formes — la langue permet de paraphraser — et à une même forme peuvent correspondre plusieurs sens — les énoncés peuvent être ambigus).
- 2) La langue doit se décrire au moyen d'un modèle fonctionnel formel qui simule cette correspondance.
- 3) Cette correspondance étant trop complexe, il faut distinguer des niveaux intermédiaires de représentation et, donc, de correspondance.

1. Cet article contient en plus une bibliographie (non-exhaustive, mais assez complète) de publications traitant de divers sujets dans le cadre de la TST.

La TST se distingue donc des théories linguistiques dominantes d'au moins deux façons. Premièrement, un modèle Sens-Texte n'est pas génératif, mais transductif, c'est-à-dire qu'il doit prendre en entrée une représentation (sémantique, phonologique ou intermédiaire) pour produire les représentations qui lui correspondent (*cf.* Kahane 2000b). Deuxièmement, alors que la plupart des théories se contentent d'un, deux, voire trois niveaux de représentation linguistique, la TST utilise sept niveaux de représentation différents: la représentation sémantique (RSém), la représentation syntaxique profonde (RSyntP), la représentation syntaxique de surface (RSyntS), la représentation morphologique profonde (RMorphP), la représentation morphologique de surface (RMorphS), la représentation phonologique profonde (RPhonP) et la représentation phonologique de surface (RPhonS).

La correspondance entre les représentations est assurée par six modules (c'est-à-dire des ensembles de règles), nommés d'après le niveau le plus profond qu'ils mettent en jeu. Les correspondances entre les sens de la langue et les concepts de la psyché, ou encore entre les phones et les sons qui les réalisent, ne font pas partie de la langue au sens strict du terme. L'étude de la réalisation des sons relève de la phonétique articulatoire (considérée dans le cadre de cette théorie comme externe à la linguistique dans son acception la plus étroite), alors que l'étude de la pensée appartient aux sciences cognitives. Un modèle Sens-Texte ne doit donc décrire que les correspondances qui existent entre les sens de la langue et les phones qui les expriment.

La Figure 1, p. 37, présente de façon schématique les niveaux de représentation de la TST et les modules qui les mettent en correspondance. Les niveaux qui nous intéressent dans cette thèse sont les niveaux sémantique et syntaxique profond et de surface. Nous en discuterons plus en détail et nous esquisserons une description des autres niveaux avant d'aborder la question des modules qui établissent les correspondances entre ces représentations.

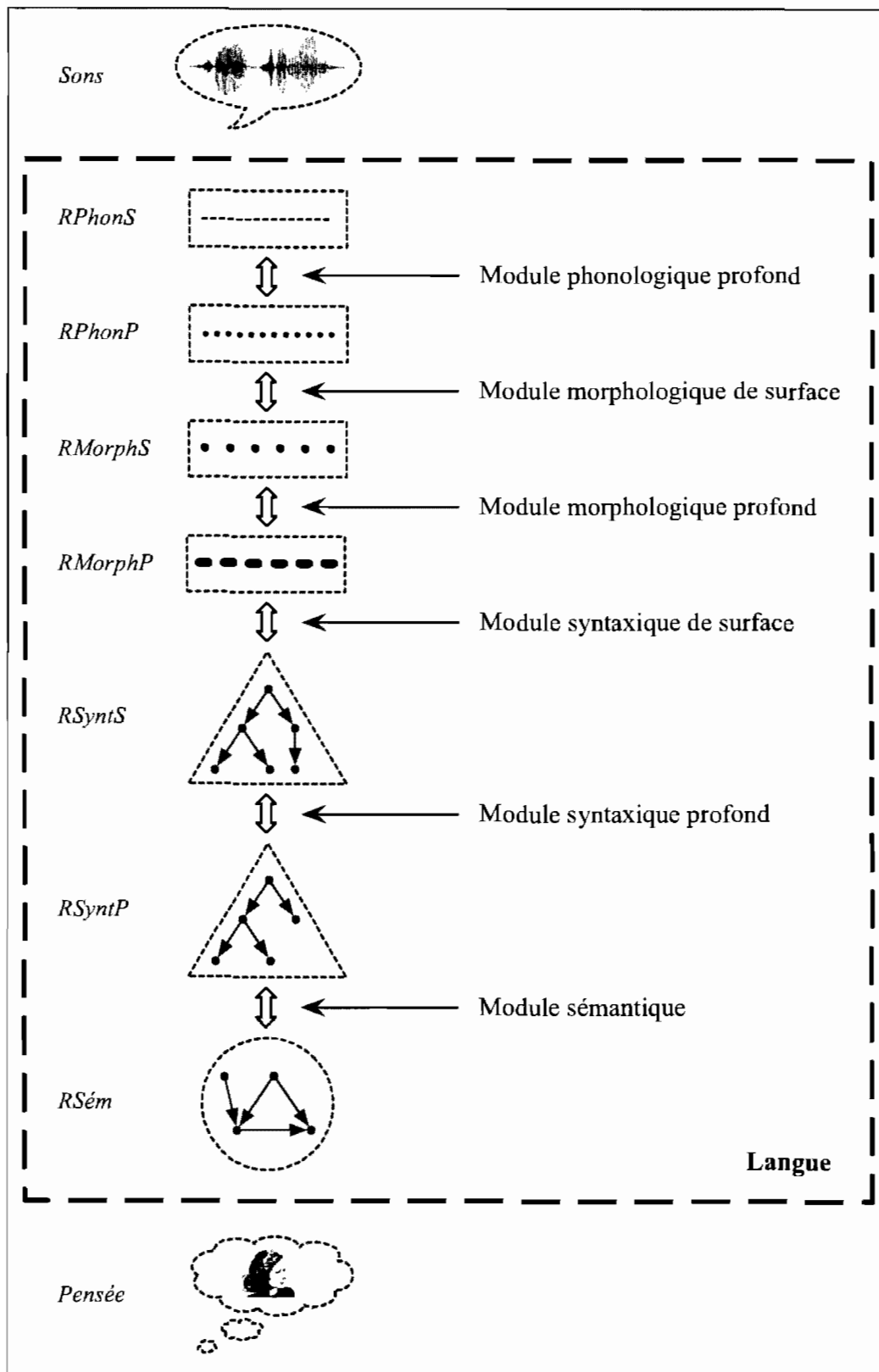


Figure 1 — L'architecture d'un modèle Sens-Texte

III. La théorie Sens-Texte

2 Les niveaux de représentation

Chaque représentation linguistique de la TST est composée de plusieurs structures, dont une est la structure porteuse. Les structures porteuses sont des graphes dont la complexité va grandissante à mesure qu'on s'approche du niveau sémantique. La représentation du sens d'un énoncé est basée sur un réseau. Au niveau syntaxique, où est représentée l'organisation des mots dans la phrase, c'est un arbre de dépendances qui sert de support à la représentation. À partir du niveau morphologique profond, les représentations sont des chaînes: d'abord des chaînes de lexèmes (en RMorphP), puis de morphèmes (en RMorphS), ensuite de phonèmes (en RPhonP) et finalement de phones (en RPhonS).

Nous n'insisterons pas sur la forme des représentations linguistiques de la TST, puisqu'elles sont très semblables à celles de la grammaire d'unification Sens-Texte, qui en sont fortement inspirées et que nous présenterons au Chapitre VII. Ce qui nous intéresse ici est surtout de donner au lecteur une idée générale de l'architecture d'un modèle Sens-Texte. Pour chaque niveau de représentation, nous montrerons en exemple la structure porteuse pour la phrase ci-dessous:

(7) *Je pense à ma sœur.*

2.1 La représentation sémantique

La RSém d'un énoncé en représente le sens. Une RSém complète est constituée de quatre structures: la structure sémantique (SSém), la structure sémantico-communicative (SSém-Comm), la structure rhétorique (SRhét) et la structure référentielle (SRéf).

La SSém est la structure porteuse d'une RSém. Elle représente le sens propositionnel exprimé par un énoncé et l'ensemble de ses paraphrases. Il ne s'agit pas d'une représentation du monde, ni même de la connaissance qu'en a le locuteur, mais bien d'une représentation du sens strictement linguistique. Formellement, la SSém est un réseau de sémantèmes liés par des relations prédicat-argument. Par

exemple, la structure sémantique simplifiée² de la phrase (7) ci-dessus est la suivante :

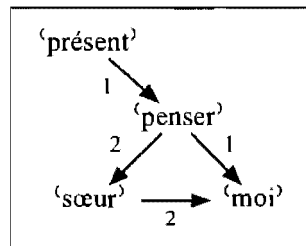


Figure 2 — La SSém de (7)

Les arcs de la SSém, qui représentent les relations prédicat-argument entre les sémantèmes, sont étiquetés par des numéros plutôt que des noms de « rôles sémantiques » comme *agent*, *patient*, etc. Les arguments d'un prédicat sont numérotés en ordre croissant d'oblicité (cf. Mel'čuk 2004).

Les sémantèmes varient d'une langue à l'autre et ne correspondent pas forcément à des unités lexicales. Au niveau sémantique, aucune distinction n'est faite entre les sens qui s'expriment par des moyens grammaticaux et ceux qui s'expriment par des lexies. De même, les parties du discours n'existent pas à ce niveau de représentation. Le sémantème 'penser', par exemple, peut aussi bien se réaliser par le verbe PENSER que par le nom PENSÉE.

La SSém ci-dessus représente non seulement le sens de (7), mais aussi de toutes ses paraphrases plus ou moins exactes. Par exemple, toutes les phrases ci-dessous (et d'autres encore) ont la même SSém puisqu'elles dénotent exactement la même situation, c'est-à-dire que leur sens propositionnel est le même.

- (8) *Je pense à ma frangine.*
- (9) *C'est à ma sœur que je pense.*
- (10) *Moi, je pense à ma sœur.*
- (11) *Moi, ma sœur, j'y pense.*
- (12) *J'y pense, à ma sœur, moi.*

2. Nous omettons dans cette figure les sens grammaticaux, sauf celui du présent de l'indicatif, que nous représentons ici par un seul nœud. Nous verrons au Chapitre VIII (§ 1.3, p. 286) que sa représentation est en fait plus complexe.

- (13) *Je pense à ma sœur.* [avec emphase]
 (14) *J'ai une pensée pour la sœur à bibi.*

C'est principalement dans la SSém-Comm que se représente la différence entre ces phrases. Cette structure encode l'organisation communicative des sens d'un énoncé, c'est-à-dire la façon dont le locuteur a choisi de les présenter. Par exemple, en (10) le sémantème 'moi' est communicativement plus saillant qu'en (12), où il est quand même plus saillant qu'en (13). Cette organisation se représente par une division de la SSém en aires communicatives marquées des huit oppositions suivantes, dont les valeurs sont mutuellement exclusives :

- 1) La thématité: rhème / thème / spécifieur.
- 2) La nouveauté: 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'articulation: unitaire / articulé.
- 8) La locutionnalité: communiqué / signalé / performé.

Nous n'allons pas décrire chacune de ces oppositions, puisque nous ne les utilisons pas dans cette thèse. Nous renvoyons le lecteur à Mel'čuk (2001c).

Les deux autres structures de la RSém ont été beaucoup moins étudiées. La SRhét encode l'information stylistique pertinente à la production d'un énoncé. C'est dans cette structure, par exemple, que s'encode la différence entre (7) et (8), qui est de l'ordre du registre. La SRéf, quant à elle, lie les sémantèmes (ou des configurations de sémantèmes) de la SSém à leur référent dans le monde réel.

2.2 La représentation syntaxique profonde

La RSyntP représente l'organisation des lexies pleines au sein de la phrase. Les lexies pleines sont celles choisies librement par le locuteur afin d'exprimer un sens. Cela exclut notamment les articles, auxiliaires, prépositions régies, etc. La

RSyntP est constituée de quatre structures: la structure syntaxique profonde (SSyntP), la structure syntaxico-communicative profonde (SSynt-CommP), la structure syntaxico-prosodique profonde (SSynt-ProsP) et la structure anaphorique profonde (SAnaphP).

La SSyntP est la structure porteuse de la RSyntP. Il s'agit d'un arbre de dépendances non ordonné; alors que la plupart des théories syntaxiques dominantes sont basées sur des structures de constituants, la TST ne fait usage que de la dépendance (*cf.* Mel'čuk 1988a, 2001a ou Mel'čuk & Polguère 2008). Les nœuds de la SSyntP sont étiquetés de noms de lexies pleines. Ces lexies peuvent être accompagnées de grammèmes profonds (nous reviendrons au Chapitre IV sur la notion de grammème). Les relations syntaxiques profondes entre les lexies sont de quatre types :

- Les relations actantielles, numérotées de I à VI, qui lient les lexies à leurs actants syntaxiques.
- Les relations attributives, $ATTR_{\text{restr(ictive)}}$ et $ATTR_{\text{qual(ificative)}}$, qui lient les modifieurs et les relatives à leur gouverneur.
- La relation appenditive, APPEND, qui rattache au verbe principal des éléments extra-structurels comme les apostrophes, les interjections, etc.
- Les relations de coordination, COORD et QUASI-COORD.

Par exemple, la SSyntP de la phrase (7) est la suivante:

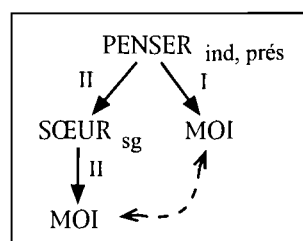


Figure 3 — La SSyntP de (7)

L'arc en pointillés entre les deux nœuds MOI appartient à la SAnaphP. Il s'agit d'un lien de coréférence. La SAnaphP ne contient pas d'autre information. La SSynt-CommP est le reflet, au niveau syntaxique profond, de la SSém-Comm. Quant à la SSynt-ProsP, elle consiste en des instructions indiquant une prosodie

porteuse de sens à appliquer à la phrase (par exemple, ironique, déclarative, interrogative, etc.).

2.3 La représentation syntaxique de surface

La RSyntS est structurellement très similaire à la RSyntP. Toutes les structures de la RSyntP ont leur équivalent en RSyntS: la structure syntaxique de surface (SSyntS), la structure syntaxico-communicative de surface (SSynt-CommS), la structure syntaxico-prosodique de surface (SSynt-ProsS) et la structure anaphorique de surface (SAnaphS).

La SSyntS est aussi un arbre de dépendances non ordonné. Elle se distingue de la SSyntP par le fait que tous les lexèmes de la phrase y sont présents, ainsi que par la nature des relations syntaxiques qui y apparaissent. Les relations syntaxiques de surface varient d'une langue à l'autre (*cf.* Mel'čuk & Pertsov 1987 pour l'anglais et Iordanskaja & Mel'čuk, à paraître, pour le français). La SSyntS pour la phrase (7) est la suivante:

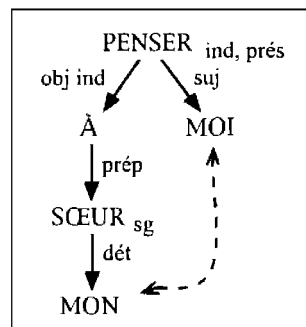


Figure 4 — La SSyntS de (7)

Les autres structures de la RSyntS sont identiques à celles de la RSyntP.

2.4 Les représentations morphologiques et phonologiques

À partir du niveau morphologique profond, les structures porteuses sont linéaires. La structure morphologique profonde (SMorphP) est une chaîne de lexèmes accompagnés de toutes les informations grammaticales nécessaires au calcul de leur forme fléchie. Par exemple, la SMorphP de (7) est la suivante:

MOI _{subj}	PENSER _{ind, prés}	À	MON _{sg, fém}	SCEUR _{sg}
----------------------	------------------------------	---	-------------------------	----------------------

Figure 5 — La SMorphP de (7)

La structure morphologique de surface (SMorphS) est une chaîne de morphèmes. Pour la phrase (7), la SMorphS est :

{MOI} ⊕ {SUBJ}	{PENS-(er)} ⊕ {IND.PRÉS}	{À}	{M-(on)} ⊕ {SG.FÉM}	{SCEUR} ⊕ {SG}
----------------	--------------------------	-----	---------------------	----------------

Figure 6 — La SMorphS de (7)

La structure phonologique profonde (SPhonP) est une chaîne de phonèmes (c'est une transcription phonologique). La SPhonP de la phrase (7) est :

/ʒə.pās.a.ma.sœr/

Figure 7 — La SPhonP de (7)

Enfin, la structure phonologique de surface (SPhonS) est une chaîne de phones (c'est une transcription phonétique, dont le degré de précision peut varier). La SPhonS de la phrase (7) est la suivante³ :

['ʃpās.a.ma.'sœ ^g :ʁ]

Figure 8 — La SPhonS de (7)

À chaque niveau entre la RSyntP et la RPhonS, s'ajoute à la structure porteuse une structure prosodique. Cette structure se raffine à mesure qu'on approche de la surface.

3 Les modules de correspondance

La transition entre les représentations se fait par l'application de règles de correspondances, regroupées dans six modules à l'interface des sept niveaux de représentation (*cf.* Figure 1, p. 37). Toutes les règles de ces modules sont composées de trois parties: deux fragments de graphes de niveaux adjacents à mettre en correspondance et un ensemble de conditions. Par exemple, la règle suivante du module

3. Ici, nous représentons notre propre prononciation.

syntaxique de surface met en correspondance un fragment de la SSyntS et un fragment de la SMorphP.

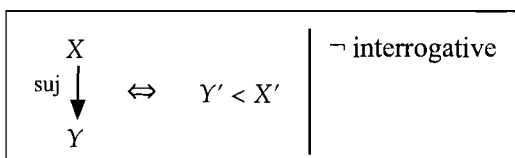


Figure 9 — Une règle du module syntaxique de surface

Cette règle modélise le fait qu'un sujet peut se placer avant son gouverneur, à condition de ne pas être dans une construction interrogative (à inversion du sujet).

Nous n'allons pas présenter en détail la forme des règles d'un modèle Sens-Texte, puisque nous y reviendrons au Chapitre VII. Ce qui nous intéresse ici est de voir quels sont les phénomènes linguistiques modélisés par chaque module (en nous plaçant dans le sens de la synthèse, et non de l'analyse).

Le module sémantique sert principalement à organiser l'énoncé en une structure syntaxiquement valide. L'arborisation de la SSém vers la SSyntP est fortement conditionnée par la SSém-Comm, puisque la SSyntP correspond plus ou moins à la hiérarchie communicative entre les sémantèmes établie par la SSém-Comm (*cf.* Polguère 1990). C'est aussi dans le module sémantique qu'a lieu la lexicalisation profonde, c'est-à-dire le choix des lexies qui expriment les sens de la SSém. Dans la même étape, on choisit les grammèmes qui expriment les «sens grammaticaux».

Le module syntaxique profond opère la lexicalisation de surface. C'est ici que sont introduits les prépositions régies et les marqueurs analytiques des grammèmes de la SSyntP, tels les articles, auxiliaires, etc. C'est aussi dans cette étape que sont choisis les collocatifs (qui, dans la RSyntP, sont représentés par des étiquettes généralisées, les fonctions lexicales — *cf.* Mel'čuk 1996b). L'organisation syntaxique est aussi raffinée dans cette transition.

C'est dans la transition de la RSyntS vers la RMorphP qu'a lieu la linéarisation. C'est aussi dans cette transition que s'opère l'accord.

Enfin, les phénomènes morphologiques sont traités dans les modules morphologiques profond et de surface, alors que les phénomènes phonologiques sont traités dans le module phonologique profond.

4 Synthèse

En bref, la TST offre un cadre intéressant parce qu'elle couvre l'ensemble des phénomènes linguistiques structuraux. Son architecture hautement stratifiée facilite l'écriture des règles des modèles linguistiques puisque le problème de la description de la langue est découpé en sous-problèmes. Un modèle Sens-Texte est un modèle fonctionnel qui met en correspondance des (représentations de) sens et des (représentations de) «textes» au sens large — écrits ou oraux. Cette correspondance se décrit à travers six modules : les modules sémantique, syntaxique profond, syntaxique de surface, morphologique profond, morphologique de surface et phonologique profond.

Dans le cadre de cette thèse, nous avons choisi de laisser de côté tous les problèmes de nature strictement morphologique. Nous visons la production de représentations syntaxiques de surface bien formées, où apparaissent tous les grammèmes nécessaires au calcul des formes fléchies. Les niveaux de représentation pertinents pour nous sont donc les niveaux sémantique, syntaxique profond et syntaxique de surface.

IV. L'étude du grammème : théorie et méthodologie

Pour un linguiste qui se propose de décrire la conjugaison française, le travail consiste essentiellement à étudier les différentes formes verbales pour isoler les signes grammaticaux qu'elles portent et en décrire les composantes (signifié, signifiant et combinatoire). Les notions de grammème et de catégorie flexionnelle sont donc au cœur de cette étude. Pourtant, s'il y a littéralement des milliers de publications sur le temps et l'aspect, assez peu d'auteurs se donnent la peine de discuter la nature des signes en jeu dans la flexion verbale. Il devient alors difficile de comprendre comment on arrive à une telle variété dans les descriptions. Par conséquent, nous allons d'abord proposer une brève caractérisation des notions fondamentales de grammème et de catégorie flexionnelle, en nous basant principalement sur les travaux de Mel'čuk, de Kahane et de Martinet. En particulier, nous verrons que la conception traditionnelle du grammème correspond à deux choses distinctes : les grammèmes à proprement parler et les grammies, c'est-à-dire les acceptions de ces grammèmes. Nous ferons également la distinction entre grammèmes profonds et grammèmes superficiels. Nous détaillerons ensuite notre méthodologie pour l'étude de la flexion verbale, dont l'originalité réside dans le fait qu'elle s'inspire de la lexicographie explicative et combinatoire pratiquée dans le cadre de la théorie Sens-Texte (Mel'čuk, Clas & Polguère 1995, Polguère 2008, Mel'čuk 2006). L'analogie entre les deux domaines d'étude est double : la division des grammèmes en grammies rappelle celle des vocables en lexies et le regroupement des grammèmes en catégories flexionnelles sur la double base de leur combinatoire et de leur sens est en certains points similaire au regroupement des vocables en parties du discours et en champs lexicaux.

1 Les notions de grammème et de grammie

Mel'čuk (1993) définit le **grammème** comme une signification qui fait partie d'une catégorie flexionnelle, une catégorie étant définie dans le même ouvrage comme «un ensemble maximum de significations qui s'excluent mutuellement

dans la même position (sémantique ou logique)» (p. 261). Le problème est que le concept de «signification», pourtant central ici, n'est défini nulle part dans cet ouvrage; il a volontairement été laissé vague. On ne sait donc pas s'il s'agit d'une entité de nature sémantique, syntaxique ou morphologique, ou encore d'une correspondance entre de telles entités¹.

Cette définition vague cherche à résoudre ce qu'on pourrait appeler le paradoxe de la polysémie des grammèmes. Les grammèmes ont en effet tendance à porter plusieurs sens. Par exemple, la forme verbale du futur simple en espagnol peut exprimer aussi bien un sens temporel [*Vendré sobre las 6* 'je viendrai vers 6 heures'] qu'un sens modal [(en réponse à ¿*Qué hora es?* 'quelle heure est-il?') *Serán las 2* 'je dirais qu'il est 2 heures' — littéralement, 'il sera 2 heures']. Dans le premier cas, le futur simple situe un événement à un moment postérieur au moment d'énonciation. Dans le second cas, par contre, il ne s'agit aucunement d'un sens temporel. La situation décrite par le verbe n'est pas présentée comme postérieure au moment d'énonciation mais plutôt comme une hypothèse sur une situation actuelle. Il y a là deux signes bien distincts puisque leurs sens, à première vue, n'ont pas grand-chose en commun. En se basant uniquement sur des critères sémantiques, on devrait classer le premier parmi les temps et le second avec les modes. On voudrait donc dire qu'il existe deux grammèmes du futur en espagnol. Pourtant, aux niveaux syntaxique, morphologique et phonologique, ces deux signes se comportent de la même façon, si bien qu'il est impossible de les distinguer. On serait donc tenté de dire qu'il s'agit de la même entité, qu'il n'y a qu'un seul futur dans cette langue.

C'est probablement ce qui a poussé Mel'čuk à affirmer que le grammème est «un faisceau de correspondances entre un ensemble de marqueurs donnés et un ensemble de sens donnés» (1993: p. 278). Autrement dit, les grammèmes servent de pivot entre des sens et des formes. Il donne l'exemple du pluriel des noms en

1. Pour Polguère (2008), une signification est «un contenu sémantique, non en tant que propriété d'une entité linguistique (comme dans le cas de sens), mais en tant qu'information exprimable par une entité linguistique» (p. 124).

anglais, que nous reprenons sous une forme légèrement modifiée ici. Le pluriel en anglais s'exprime, sauf exceptions, par /s/, /z/ ou /əz/. Il peut porter un certain nombre de sens, par exemple 'plus d'un X' [*Cats were sleeping there*] ou 'tous les X' [*Piranhas are dangerous*]². Il est bien entendu possible d'établir une correspondance entre chacun des sens et chacune des formes, comme dans la Figure 10a, mais alors on a l'impression de passer à côté d'une généralisation importante. Un modèle du type de ce qu'on trouve à la Figure 10b semble plus élégant et surtout plus plausible d'un point de vue cognitif (cette figure est adaptée de Mel'čuk, *ibid.*).

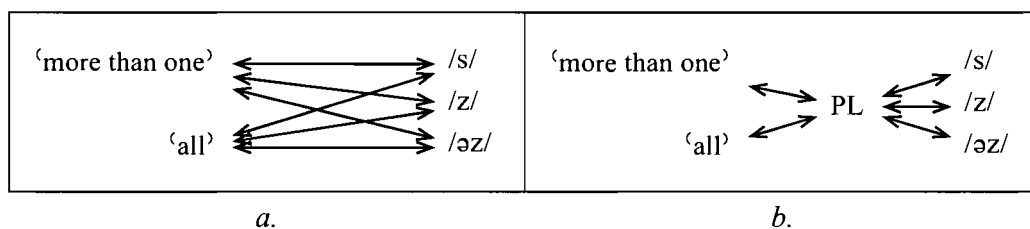


Figure 10 — Le grammème comme intermédiaire entre sens et formes

En ce sens, Mel'čuk conçoit donc les grammèmes (ou du moins, certains grammèmes) comme des objets des représentations syntaxiques (profonde ou de surface) qui servent d'intermédiaire entre des entités sémantiques (les sens grammaticaux) d'un côté, et de l'autre des entités syntaxiques de surface (les auxiliaires, articles et autres mots grammaticaux) ou morphologiques (les morphèmes flexionnels). Il s'agit donc pour lui d'une unité d'ordre sub-sémiotique, c'est-à-dire que le grammème n'est pas un signe linguistique, mais plutôt une composante de signe. Cela résout le problème de la polysémie des grammèmes en évitant d'une part de dire qu'à chaque sens distinct correspond un grammème différent et d'autre part d'affirmer qu'un grammème est un signe qui peut porter plusieurs sens (ce qui contredirait la définition du signe linguistique donnée dans le même ouvrage).

Pourtant, il est évident qu'il existe des signes grammaticaux correspondant à ce qu'on appelle «grammème». Les correspondances entre les sens grammaticaux et

2. Ces descriptions sémantiques ne sont évidemment que très approximatives. Ce n'est pas notre propos ici de décrire la sémantique de ces grammèmes et nous renvoyons le lecteur intéressé par la question à Beyssade & Dobrovie-Sorin (2005).

les grammèmes qui les expriment sont bien des signes : ils ont un sens, une forme et une combinatoire clairement identifiables. On retrouve l'équivalent dans le domaine lexical. Les différentes acceptions d'un vocable polysémique peuvent avoir des sens très éloignés. Par exemple, le vocable **FACE** regroupe une lexie qui désigne le visage et une autre qui désigne une des surfaces qui composent un objet. Ces deux signes ont des sens assez distincts pour qu'on ne puisse pas les regrouper dans un même champ sémantique. La première lexie appartient aux champs sémantiques des parties du corps ou encore des sentiments (puisque le visage peut servir à communiquer les émotions), alors que la seconde appartient plutôt, par exemple, au champ sémantique de la géométrie. Pourtant, on regroupe ces deux lexies sous un même vocable puisqu'ils partagent des caractéristiques non triviales à tous les niveaux de fonctionnement de la langue (y compris au niveau sémantique). Les signes grammaticaux montrent le même genre d'organisation. Les signes qui partagent des caractéristiques non triviales sont regroupés sous une unité d'un niveau d'abstraction plus élevé, équivalent à celui du vocable, que nous appelons **grammème**. Cette entité abstraite regroupe des signes grammaticaux similaires. Nous utiliserons le terme **grammie** pour faire référence aux différentes acceptions d'un grammème. Le terme vient de Kahane (2002), pour qui les grammies sont des signes profonds au même titre que les lexies, c'est-à-dire des (regroupement de) signes dont le signifié est un élément de la structure sémantique et dont le signifiant est un élément de la structure syntaxique. Ce qui différencie les grammies des lexies dans ce cadre est la nature de leur signifiant : les grammies ont comme signifiant un grammème, alors que les lexies ont comme signifiant un lexème³. Ainsi, pour reprendre l'exemple donné plus haut, il y aurait en espagnol au moins deux grammies, FUTUR₁ et FUTUR₂, qui seraient deux acceptions du grammème **futur**.

Dans le cadre de la TST classique, les grammèmes apparaissent aux niveaux syntaxiques profond et superficiel. Cependant, ce ne sont pas tous les grammèmes qui peuvent apparaître aux deux niveaux de représentation. Par exemple, la déter-

3. Ce que Kahane appelle *lexème* correspond à peu près à ce que la lexicologie explicative et combinatoire désigne habituellement par le terme *vocable*.

mination en français est toujours exprimée par un article. Le signifiant étant un lexème, il doit avoir son propre nœud au niveau syntaxique de surface. Il n'y a donc de grammèmes **défini** et **indéfini** en français qu'au niveau syntaxique profond, puisqu'en surface ils sont déjà exprimés par les lexèmes LE et UN. Ainsi, il est utile de distinguer deux types de grammèmes: les **grammèmes superficiels**, qui n'apparaissent qu'en syntaxe de surface et qui sont mis en correspondance avec des éléments de la représentation morphologique profonde (des morphèmes, des prosodèmes ou des relations de précédence entre mots-formes), ainsi que les **grammèmes profonds**, qui apparaissent en syntaxe profonde et qui doivent être mis en correspondance avec des éléments de la représentation syntaxique de surface (des lexèmes ou des grammèmes superficiels). Dans le cadre de GUST, comme nous le verrons au Chapitre VII, il n'y a pas de niveau syntaxique profond. Les grammèmes de GUST sont donc des grammèmes superficiels. Inversement, Mel'čuk traite principalement des grammèmes profonds dans son *Cours de morphologie générale*.

Cette distinction faite entre grammème (profond ou superficiel) et grammie, revenons sur la définition que donne Mel'čuk (1993) du grammème. Dans cet ouvrage, il le définit comme un des éléments d'une **catégorie flexionnelle**. Cette dernière notion est donc plus élémentaire selon lui. En effet, il s'agit d'une caractéristique importante du grammème que de faire partie d'une catégorie. Les grammèmes ne sont des grammèmes que s'ils sont en opposition avec d'autres grammèmes, avec lesquels ils forment un ensemble dont les éléments sont mutuellement exclusifs. Ce qui fait qu'une catégorie est une catégorie flexionnelle, c'est la nature justement flexionnelle de ses éléments. Passons en revue quelques propriétés qui distinguent la flexion⁴ de la dérivation.

4. Nous utilisons ce terme dans son sens le plus large et considérons la flexion comme un phénomène qui n'est pas seulement morphologique, ce qui inclut les auxiliaires.

2 Flexion vs dérivation : propriétés distinctives

De toutes les propriétés évoquées par les linguistes pour caractériser la flexion (par opposition à la dérivation), son caractère obligatoire est sans doute la plus centrale et celle sur laquelle il y a le plus large consensus. Jakobson (1959) disait, en commentant les travaux de Boas (1938) : «[...] la vraie différence entre les langues ne réside pas dans ce qu'elles peuvent ou ne peuvent pas exprimer, mais dans ce que les locuteurs doivent ou ne doivent pas transmettre». De cette propriété en découle d'autres qui sont souvent utilisées pour définir la flexion. Par exemple, Riegel, Pellat et Rioul (1994) précisent que les morphèmes grammaticaux appartiennent à des classes en général moins nombreuses que les morphèmes lexicaux. Or, c'est une conséquence prévisible du caractère obligatoire de la flexion. On aurait en effet du mal à imaginer, par exemple, une langue où il y aurait 300 modes et 500 temps pour les verbes. Il découle également de ce caractère obligatoire le fait que les sens grammaticaux tendent à être plus abstraits ou plus vagues que les sens lexicaux. Puisque les signes flexionnels doivent pouvoir se combiner à toutes les lexies d'une même classe ils ne peuvent pas, en général, porter de sens très concret ou précis puisque les lexies auxquelles ils doivent se combiner ne partagent pas forcément de composante sémantique commune⁵.

Mel'čuk (1993) donne encore six propriétés qui permettent de distinguer les grammèmes des dérivatèmes :

- 1) Ils résistent mieux à la phraséologisation.
- 2) Ils ont normalement une combinatoire plus large.
- 3) Ils ont tendance à avoir une expression plus régulière.
- 4) Seuls les grammèmes figurent dans les règles syntaxiques d'accord ou de concordance de temps.
- 5) Ils ont tendance à être exprimés plus loin de la racine.
- 6) Ils ne modifient pas la partie du discours du radical⁶ auquel ils s'attachent.

5. Mel'čuk (1993) rappelle qu'il existe des cas limites. Il cite en exemple une analyse de Sapir (1915) selon laquelle en Nootka les verbes exprimeraient des caractéristiques physiques de leur sujet par le biais d'un préfixe flexionnel. Les sens des grammèmes de la catégorie flexionnelle en jeu sont précis et concrets : 'borgne', 'bossu', 'trop gros', etc.

Mentionnons également la propriété suivante, relevée par Riegel, Pellat et Rioul (1994):

- 7) Le système flexionnel résiste au changement. L'introduction d'un nouveau grammème constituerait une modification importante dans la langue en général.

Une fois qu'on a identifié dans la langue les signes flexionnels (par opposition aux signes dérivationnels) grâce à ces propriétés, il reste à les regrouper en catégories flexionnelles. Nous proposons dans les pages qui suivent une méthodologie pour ce faire.

3 Une approche discrète en deux étapes

Comme nous l'avons déjà mentionné plus haut, un grammème peut très bien correspondre à plusieurs sens différents. En fait, un des principaux problèmes que pose l'étude des grammèmes est justement leur polysémie, souvent riche et subtile. Il est difficile de dégager un système cohérent de la conjugaison française si on tente de décrire à la fois tous les sens d'un même grammème verbal. Pour faire un parallèle avec la lexicographie, cela reviendrait à tenter de décrire le sens d'un vocable sans faire la distinction entre les lexies qui le constituent. Nous prôtons une approche résolument discrète (en ce sens que nous croyons qu'il est possible d'isoler les différentes acceptions d'un grammème) et en deux étapes. Dans un premier temps, nous ne considérons que la grammie de base de chaque grammème, c'est-à-dire son acception propre. Ce n'est qu'une fois que nous avons dégagé un système cohérent à partir de ces grammies de base que nous prenons en considération les autres acceptions des grammèmes.

6. Rappelons que la racine d'un mot-forme est la lexie à partir de laquelle est construit ce mot-forme, alors que le radical auquel s'attache un affixe peut contenir, en plus de la racine, un certain nombre d'affixes.

4 La grammie de base d'un grammème

Lorsqu'on examine les différents sens que peut exprimer un grammème, on perçoit intuitivement qu'il y en a un qui est plus saillant que les autres. Les grammèmes ont, comme les vocables, un sens de base et des sens figurés ou dérivés. L'idée n'a rien de neuf, elle se trouve formulée presque telle quelle dans la grammaire espagnole de Bello, au XIX^e siècle :

Il faut d'abord noter que chaque forme verbale a habituellement, en plus de sa valeur propre et fondamentale, d'autres valeurs différentes dans lesquelles se transforme la première selon certaines règles générales. Ainsi, nous distinguons parmi les formes d'un verbe un sens fondamental à partir duquel sont dérivés deux autres sens : le sens secondaire et le sens métaphorique.

Bello (1982 [1847]: p. 200), notre traduction

Nous appellerons **grammie de base** l'acception d'un grammème qui correspond à son «sens propre», par analogie avec les lexies de base de la lexicologie explicative et combinatoire (Mel'čuk, Clas & Polguère 1995). Il est en général évident, de façon intuitive, quelle est l'acception de base d'un grammème. Pourtant, nous ne pouvons pas formuler de critères parfaitement clairs qui permettent d'identifier systématiquement la grammie de base de tous les grammèmes. Nous proposons ici quelques pistes et discutons des problèmes liés à chaque critère.

Il y a logiquement deux types de critères imaginables : ceux basés sur le sens des grammies et ceux qui s'appuient sur leur combinatoire. Il n'est pas possible d'avoir des critères basés sur la forme des grammies puisque toutes les acceptions d'un même grammème sont, par définition, associées aux mêmes formes.

4.1 Les critères basés sur le sens

Si on demandait à brûle-pourpoint à un locuteur natif du français quelle est la signification de *il mangera*, on s'attendrait à ce qu'il nous réponde que c'est une activité qui se déroule à un moment ultérieur. Nous croyons pouvoir affirmer sans trop nous tromper qu'il y a un certain consensus parmi les francophones pour dire que le futur, pris hors contexte, sert d'abord à situer un fait dans l'avenir. Après, si

on y réfléchit un peu plus et qu'on met le verbe dans divers contextes, on arrivera peut-être à identifier d'autres sens qui peuvent s'exprimer par le futur mais ces autres acceptions du grammème futur ne viennent à l'esprit que si le contexte le suggère. Le sens qui est perçu intuitivement par les locuteurs comme le «sens premier» correspond à ce que nous appelons la grammaire de base du grammème. De façon triviale, le premier critère que nous pourrions formuler, celui qui capte le mieux l'essence de ce qu'est la grammaire de base d'un grammème, serait donc le suivant :

Critère de l'interprétation spontanée

C'est l'interprétation la plus courante, celle qui vient spontanément à l'esprit des locuteurs hors contexte, qui correspond à la grammaire de base d'un grammème.

Il est évident que ce critère ouvre la porte à une certaine subjectivité et doit être utilisé avec prudence. Il n'en demeure pas moins que c'est, de tous les critères que nous pouvons formuler, celui qui correspond le mieux à ce que nous tentons de définir comme «grammaire de base».

Ce critère n'est toutefois pas toujours applicable. Bien entendu, il ne l'est pas facilement dans le cas des grammèmes qui ne sont associés à aucun sens (comme les grammèmes d'accord). Mais encore, même quand les grammèmes sont associés à des sens clairement identifiables, la situation se trouve compliquée par le fait que certains grammèmes sont «marqués»⁷ alors que d'autres ne le sont pas. Par exemple, dans le paradigme des temps verbaux en français, on peut dire que le futur est marqué alors que le présent ne l'est pas. Le sens de base des grammèmes marqués étant plus saillant que celui des grammèmes non marqués, il est tout de suite accessible au locuteur, alors que celui des grammèmes non marqués peut être plus difficile à trouver. Ainsi, si on demande à un francophone ce que *je mangerai* signifie, on s'attend à ce qu'il précise que l'activité en question se déroule dans le

7. La distinction «marqué» ~ «non marqué», depuis son introduction par Jakobson et le Cercle linguistique de Prague, n'a jamais, à notre connaissance, été précisément définie. Mel'čuk (1994: pp. 11-12) donne un certain nombre de propriétés qui caractérisent ces notions.

futur. Par contre, si on lui demande ce que signifie *je mange*, il se lancera vraisemblablement dans une description de ce que ‘manger’ signifie, et non de ce que le temps présent signifie. Il faut donc chercher à combiner ce critère avec d’autres qui puissent combler ses lacunes.

De façon empirique, on peut observer que, comme c’est le cas pour la lexie de base d’un vocable, le sens de base d’un grammème est souvent inclus dans celui des autres acceptions de ce même grammème. Notamment, on peut trouver des liens métaphoriques entre les sens dérivés et le sens de base. Par exemple, le lien sémantique entre un présent qui désigne réellement le moment d’énonciation [*Qu’est-ce que tu fais?*] et un présent qui désigne un futur rapproché [*Le train part à 8h34 demain*] peut se décrire par une métaphore⁸:

- ‘X-présent₁’ = ‘X a lieu maintenant’
- ‘X-présent₂’ = ‘X est très rapproché dans le futur ou inexorable [tellement que c’est comme si X-présent₁]’.

On peut donc formuler un deuxième critère à partir de cette observation :

Critère de l’inclusion de sens

Si les grammies A et B correspondent au même grammème G et que le sens de A est inclus dans le sens de B (soit par simple inclusion — directe ou indirecte — soit par un lien métaphorique), alors A est la grammie de base de G.

Mel’čuk, Clas & Polguère (1995) utilisent le même critère dans le domaine lexical. Dans bien des cas, il suffit à identifier la grammie de base d’un grammème.

Il y a cependant des cas où le critère de l’inclusion de sens ne permet pas d’identifier ce qu’intuitivement nous aimerions reconnaître comme la grammie de base d’un grammème. Il y a deux cas de figure logiquement possibles où ce critère ne donne pas les résultats voulus: **1)** le critère n’est pas applicable parce que les sens des grammies d’un grammème ne s’incluent pas mutuellement, ou **2)** c’est la grammie de base qui inclut sémantiquement d’autres grammies. Nous n’avons pas

8. Nous proposons une solution différente au Chapitre VI (§ 6.2, p. 193).

trouvé d'exemple illustrant ce second cas de figure, qui constituerait un contre-exemple invalidant notre critère, mais nous en avons trouvé pour le premier.

Nous avons vu au début de ce chapitre qu'il y avait en espagnol au moins deux grammies correspondant au futur simple: une dont le sens est strictement temporel (c'est l'équivalent du FUTUR₁ français que nous venons de voir) et une autre dont le sens est plutôt modal (elle indique une hypothèse). Elles apparaissent respectivement dans les phrases (15) et (16) ci-dessous (nous reprenons ici les exemples donnés plus haut pour le bénéfice du lecteur):

(15) *Vendré sobre las 6.*
Je viendrai à 6 heures.

(16) – *¿Qué hora es? – Serán las 2.*
– Quelle heure est-il? – Je dirais qu'il est 2 heures (littéralement, *il sera 2 heures*).

Intuitivement, on perçoit le futur de la phrase (15) comme étant l'acceptation de base du grammème futur, mais notre critère ne peut pas être utilisé ici avec les résultats attendus si on adopte l'analyse sémantique suivante de ces deux grammies:

- 'X-futur₁' = 'X a lieu après maintenant'
- 'X-futur₂' = 'je suppose que X a lieu maintenant'

En effet, bien qu'il y ait intersection de sens entre les deux grammies ('maintenant'), il n'y a pas d'inclusion. Le critère de l'inclusion de sens ne peut donc pas être utilisé. Mais cela peut justement être un indice que la description sémantique des grammies en question est incorrecte. On serait alors tenté d'adopter une position semblable à celle de Riegel, Pellat & Rioul, qui affirment à propos de l'équivalent français⁹ du FUTUR₂ espagnol:

Quand un procès contemporain de l'énonciation n'est pas avéré, le futur simple est employé pour le présenter comme une hypothèse, que l'avenir confirmera ou permettra de vérifier.

Riegel, Pellat & Rioul (1994: p. 314)

9. Le futur à valeur de conjecture s'utilise beaucoup moins en français qu'en espagnol, mais on le trouve notamment dans des phrases comme – *Où est Pierre? Il devrait être arrivé.* – *Il se sera perdu en chemin.*

Une telle analyse permet de résoudre le problème de l'inclusion de sens pour cette paire de grammies :

- 'X-futur₁' = 'X a lieu après maintenant'
- 'X-futur₂' = 'je suppose que X a lieu maintenant mais je ne peux pas le vérifier maintenant [comme si X-futur₁]'

Enfin, si le critère de l'inclusion sémantique doit être utilisé avec prudence, il n'en demeure pas moins utile, et peut même guider l'analyse sémantique des grammies.

De façon générale, les critères basés sur le sens posent deux problèmes. Premièrement, les grammèmes ne sont pas tous associés à un sens. Si un grammème a plusieurs acceptions non porteuses de sens, cela pose problème pour le critère de l'inclusion sémantique. Ajoutons que les grammèmes non porteurs de sens peuvent également mettre en échec le critère de l'interprétation spontanée en rendant difficile l'identification d'un sens «intuitivement évident» pour le locuteur.

Deuxièmement, si les grammèmes de temps se prêtent plutôt bien au jeu de la décomposition sémantique, il est loin d'être évident que cela puisse être le cas de la majorité des grammèmes. En effet, beaucoup de grammèmes ne se laissent pas décrire par une paraphrase. Leur signifié peut très bien, par exemple, être un élément de la structure communicative. C'est le cas notamment du suffixe -WA en japonais, qui sert à marquer l'élément thématique de la phrase (voir § 5, p. 60, pour d'autres exemples). Il devient alors difficile d'utiliser des critères sémantiques pour de tels grammèmes. Nous proposons donc d'autres critères, basés sur la combinatoire, qui pourraient compléter les critères sémantiques.

4.2 Les critères basés sur la combinatoire

On s'attend à ce que, de façon générale, la grammie de base d'un grammème ait une combinatoire moins contraignante que les autres. Vu la définition que nous avons adoptée pour la notion de grammème, la combinatoire des grammies d'un même grammème ne peut pas différer au-delà du niveau syntaxique profond. Les

propriétés de combinatoire à considérer sont donc limitées à deux types : les configurations syntaxiques dans lesquelles le grammème peut apparaître lorsqu'il exprime tel ou tel sens et les lexies auxquelles il peut se combiner.

Tout d'abord, il semble assez naturel de vouloir exploiter, pour les grammèmes verbaux, la position privilégiée du verbe en tant que sommet syntaxique de la proposition :

Critère du sommet syntaxique (provisoire)

La grammie de base d'un grammème verbal peut être rattachée au verbe principal de la proposition.

Cependant, ce critère pose au moins deux problèmes. D'abord, il y a des grammèmes verbaux qui ne peuvent jamais apparaître dans cette position. C'est le cas par exemple des participes (présent et passé). Ensuite, pour les grammèmes auxquels il peut s'appliquer, il ne donne pas toujours les résultats voulus. Par exemple, le subjonctif n'apparaît en général que dans une subordonnée (puisqu'il sert justement à marquer la subordination, dans une certaine mesure). Il peut cependant apparaître sur le verbe principal lorsqu'il exprime un ordre à la troisième personne [*Qu'il aille demander à son chef!* ~ *Va demander à ton chef!*]. On ne voudrait pas pour autant conclure que le subjonctif sert d'abord à exprimer un ordre, puisqu'il est assez évident que c'est un emploi un peu détourné de ce grammème.

Pour remédier à ces problèmes, on peut formuler un critère plus général qui consiste à tester la flexibilité de la combinatoire syntaxique du grammème :

Critère de la polyvalence syntaxique

La grammie de base d'un grammème est celle qui peut être utilisée dans les contextes syntaxiques les plus variés.

Ainsi, s'il est vrai que le subjonctif ne peut être rattaché au sommet syntaxique que s'il exprime un ordre, il est strictement contraint à cette position lorsqu'il correspond à ce sens. Au contraire, lorsque le subjonctif marque la subordination, il peut occuper différentes positions dans la représentation syntaxique : sujet d'un verbe [*Que vous arriviez à l'heure m'importe*], complément d'un verbe [*Je veux*

que vous arriviez à l'heure], complément d'une conjonction [*À moins que vous n'arriviez en retard, tout devrait bien se passer*], etc.

Enfin, on peut imaginer un critère similaire dans le domaine lexical :

Critère de la polyvalence lexicale

La grammie de base d'un grammème est celle qui est compatible avec le plus grand nombre de lexies.

Ce critère ne s'applique que s'il y a une incompatibilité sémantique entre l'acception de base et certaines lexies. Si l'incompatibilité est syntaxique ou morphologique, c'est alors tout le grammème qui est affecté, peu importe le sens qu'il exprime. Par exemple, la combinatoire morphologique du verbe *FRIRE* ne permet pas de le combiner à l'imparfait, et ce, sans égard au sens exprimé par ce dernier.

De façon générale, les critères basés sur la combinatoire sont d'une utilité limitée puisque les différentes acceptions d'un grammème sont toutes mises en correspondance avec un seul et même objet au niveau syntaxique profond. Ce n'est donc que dans l'interface entre le niveau sémantique et le niveau syntaxique profond que peuvent s'observer les différences de combinatoire. Par conséquent, les diverses acceptions d'un même grammème tendent à avoir la même combinatoire.

Pour conclure, insistons sur le fait que ces critères ne permettent pas, de façon isolée, de repérer systématiquement la grammie de base d'un grammème. Par contre, pris ensemble, ils peuvent aider à guider le linguiste face à un cas difficile.

Identifier la grammie de base d'un grammème est une étape très importante dans notre méthodologie, puisque c'est cette acception qui sert de base à la classification des grammèmes en catégories flexionnelles. Nous porterons donc, dans notre travail, une attention toute particulière à l'acception de base de chaque grammème à l'étude, quitte à négliger ses autres acceptions dans un premier temps (nous pourrions y revenir quand nous aurons un squelette solide pour notre modèle). Une partie importante de la description d'une grammie est l'analyse de

son sens. Nous allons maintenant voir comment nous procédons à la description sémantique des grammies.

5 La description du sens grammatical

Le sens des grammies, tout comme celui des unités lexicales, doit être décrit de façon rigoureuse. Les recherches en sémantique formelle ont souvent porté sur le temps grammatical (citons par exemple Reichenbach 1947, Dowty 1972 ou encore Partee 1984), probablement parce que cette catégorie flexionnelle, en plus d'être très répandue dans les langues en général, se prête bien à la formalisation. Les formalismes varient, mais en fin de compte ils encodent pour la plupart des relations de précédence temporelle entre des points (ou des intervalles) sur l'axe du temps. Dans la majorité des théories, ces relations de précédence sont encodées comme des comparaisons entre deux ordres de grandeur.

Les représentations de ce type sont très utiles, notamment pour le traitement automatique des langues. En fait, étant donné l'état actuel des connaissances et des techniques en linguistique informatique, elles sont probablement les mieux adaptées. Cependant, nous ne visons pas seulement le traitement automatique des langues; nous cherchons à décrire les langues de la façon la plus complète et élégante possible. Pour le but que nous poursuivons, ce type de représentation ne semble pas satisfaisant.

Tout d'abord, les représentations de ce genre ne permettent pas de tenir compte de la perspective communicative qu'encodent les signes de la langue. Par exemple, Gosselin (1996) définit l'accompli par la formule $[I,II] \text{ POST } [B1,B2]$ (= l'intervalle de référence est postérieur à l'intervalle du fait en question). Le sens étant réduit à une relation de précédence temporelle entre deux points¹⁰, toute relation de précédence entre deux intervalles a toujours un équivalent converse, c'est-à-dire que $[I,II] \text{ POST } [B1,B2] \equiv [B1,B2] \text{ ANT } [I,II]$ (= l'intervalle du fait en ques-

10. Gosselin (1996) définit les relations entre intervalles par des relations entre les bornes de ces intervalles. Ainsi, $[I,II] \text{ POST } [B1,B2]$ se définit par $B2 < I$.

tion est antérieur à l'intervalle de référence)¹¹. Autrement dit, ce type de représentation n'encode pas la différence de perspective qu'il y a entre 'le fait est avant l'intervalle de référence' et 'l'intervalle de référence est après le fait'. À première vue, cela n'a aucune importance, puisque après tout il est vrai que si A est après B, alors B est avant A. Or, justement, nous croyons que cette différence de perspective est pertinente dans la langue (nous y reviendrons plus loin, Ch. VIII, § 1.3, p. 286).

D'autre part, les représentations formelles qui ne font que placer des objets sur un axe temporel n'encodent pas toujours tout le sens des différentes acceptions des grammèmes de temps, qui comporte souvent des composantes non temporelles. Par exemple, nous venons de voir que les sens dérivés d'un grammème entretiennent souvent un lien métaphorique avec d'autres acceptions de ce grammème (§ 4, p. 53). Les représentations constituées seulement de relations de précedence temporelle entre des points ne peuvent pas encoder cette information.

Bien entendu, on peut très bien étendre les formalismes de ce type pour qu'ils puissent encoder d'une façon ou d'une autre les différences de perspective communicative ou les composantes sémantiques non temporelles. Néanmoins, ces représentations ne s'intègrent pas bien dans un modèle global de la langue. Les stratégies formelles communément utilisées pour la description des sens temporels, si elles sont très bien adaptées pour cette tâche, le sont beaucoup moins pour la description de phénomènes linguistiques sans lien avec le temps. Nous cherchons à homogénéiser au maximum la description de tous les niveaux de fonctionnement de la langue; il est donc important que la description sémantique des grammies s'harmonise avec le reste du modèle linguistique.

Pour ces raisons, nous préférons utiliser des décompositions sémantiques du type de celles utilisées en lexicographie, c'est-à-dire des paraphrases, pour décrire le sens des grammèmes de temps. Bien évidemment, le sens des grammèmes ne se

11. En effet, [B1,B2] ANT [I,II] est également défini par $B2 < I$. On peut d'ailleurs se demander pourquoi Gosselin a jugé utile de créer deux relations entre intervalles, ANT et POST, si les deux se définissent précisément par la même relation entre les deux mêmes points. Peut-être cherchait-il justement à encoder une différence de perspective?

laisse pas toujours bien décrire au moyen d'une telle décomposition sémantique. Premièrement, il y a des grammèmes (ceux d'accord notamment), qui n'expriment tout simplement aucun sens. Leur signifié ne peut donc pas se décrire au moyen d'une paraphrase. Deuxièmement, il existe des grammèmes dont le signifié est d'ordre communicatif, et non sémantique (nous venons de voir l'exemple du –WA japonais; nous en verrons d'autres plus loin, Ch. VIII, § 1, p. 283). Troisièmement, il existe dans le domaine grammatical des signes équivalents aux pragmatèmes¹² dans le domaine lexical. Par exemple, les marqueurs grammaticaux de politesse dans certaines langues (le coréen et le japonais offrent de beaux spécimens) sont probablement mieux décrits par les contextes extra-linguistiques dans lesquels ils doivent s'utiliser que par une paraphrase. Enfin, ces quelques contre-exemples n'épuisent certainement pas les possibilités. Il est clair que les signifiés grammaticaux ne se laissent pas toujours décrire au moyen de paraphrases et nous ne prétendons aucunement que la paraphrase est le moyen le mieux adapté à la description sémantique des grammèmes en général. Néanmoins, les grammèmes qui nous intéressent ici, soit ceux liés au temps grammatical, se prêtent assez bien à la description par décomposition sémantique telle qu'on la pratique en lexicographie. Partout où cela est possible, nous préférons avoir recours à cette technique plutôt qu'à n'importe quelle autre.

6 Le regroupement des grammèmes en catégories flexionnelles

Les catégories flexionnelles ne sont pas directement observables. Ce qui est observable, ce sont les signes: on peut en général en percevoir la forme correspondante et en identifier le sens et la combinatoire par diverses techniques. Le travail du linguiste consiste, entre autres choses, à faire ressortir le système qui se cache derrière les signes observables.

Précisons tout d'abord que ce sont les grammèmes que nous cherchons à regrouper en catégories flexionnelles, et non les grammies. En effet, c'est à ce

12. C'est-à-dire des lexies qui sont liées au contexte d'énonciation, comme 「BON APPÉTIT」.

niveau que s'organisent les systèmes grammaticaux. Par exemple, on a vu plus haut que le grammème **futur** en espagnol correspond à au moins deux grammies (cf. § 1, p. 46). Pourtant, le paradigme dont fait partie le futur ne change pas, peu importe la grammie en jeu.

Comme nous l'avons mentionné plus haut, les grammèmes ne sont des grammèmes que s'ils font partie d'une catégorie flexionnelle. Il faut donc, une fois identifiés les signes grammaticaux en jeu, regrouper les grammèmes en un certain nombre de catégories pour en dégager un système cohérent. Alors, sur quels critères doit-on baser ces regroupements? Il existe logiquement trois types majeurs de critères possibles, correspondant aux trois composantes des signes en jeu. On peut en effet regrouper les grammèmes en catégories flexionnelles selon le sens de leur grammie de base, selon les formes auxquelles ils sont associés ou encore selon leur combinatoire. On peut bien entendu considérer des critères mixtes, basés sur plus d'une composante.

Nous ne connaissons aucun auteur qui propose de regrouper les grammèmes selon les formes qui leur sont associées. Étant donné le caractère arbitraire du signifiant des signes linguistiques, cette position nous semble intenable. En outre, les grammèmes ont tendance à être exprimés cumulativement.

6.1 Les critères basés sur la combinatoire

La combinatoire des signes grammaticaux offre une base intéressante pour formuler des critères rigoureux et vérifiables puisqu'elle est relativement facile à observer. Martinet (1979) propose une méthode, reprise notamment par Touratier (1996), essentiellement basée sur la combinatoire. Ces deux auteurs arrivent d'ailleurs à des résultats similaires, ce qui démontre le caractère reproductible de cette méthode. Elle repose principalement sur deux critères, exposés dans la *Grammaire fonctionnelle du français* de Martinet (1979), qui vont de paire et que nous reformulons ici.

Critère de l'exclusion mutuelle

Les grammèmes d'une même catégorie flexionnelle sont mutuellement exclusifs.

Critère de la similitude de combinatoire

Les grammèmes d'une même catégorie flexionnelle ont une combinatoire similaire.

Ces deux principes sont les deux facettes du test de commutation en contexte qui est très largement utilisé dans plusieurs domaines de la linguistique. Il sert par exemple à identifier les allophones d'un phonème ou encore les allomorphes d'un morphème.

Ainsi, par exemple, le masculin et le féminin des adjectifs en français correspondent à deux grammèmes d'une même catégorie flexionnelle puisque, en plus d'être mutuellement exclusifs, ils présentent la même combinatoire: les deux se combinent aux adjectifs et sont compatibles avec la catégorie flexionnelle de nombre.

Notons que la combinatoire qui unit des grammèmes au sein d'une même catégorie flexionnelle peut comprendre des éléments «négatifs». C'est ainsi que Martinet (1979) conclut que le participe et l'infinitif, par exemple, font partie de la même catégorie puisque les deux sont incompatibles avec la catégorie flexionnelle de temps (en plus d'être mutuellement exclusifs et de partager d'autres traits de combinatoire).

Cependant, les critères basés uniquement sur la combinatoire ne sont pas entièrement satisfaisants, à notre avis. Le modèle de la conjugaison française de Touratier (1996), basé exclusivement sur ces critères distributionnels, illustre bien les problèmes liés à cette méthodologie. Son raisonnement est le suivant. D'abord, l'imparfait peut se combiner avec le subjonctif (ce qui donne le subjonctif imparfait), alors que le passé simple ne le peut pas. Tout comme le subjonctif, le futur simple peut se combiner avec l'imparfait (ce qui donne le conditionnel), mais pas avec le passé. Donc, le futur fait partie de la même catégorie flexionnelle que le subjonctif, puisqu'ils ont la même combinatoire, et que le passé, puisqu'ils

s'excluent mutuellement (un verbe ne pouvant pas porter à la fois le futur simple et le passé simple). Le subjonctif ne pouvant être autre chose qu'un mode, le futur simple et le passé simple sont aussi des modes.

Ce raisonnement ne nous semble pas convaincant. Cette combinatoire en apparence capricieuse ne permet pas de tirer les conclusions que tire Touratier. Le fait que le futur ne se combine pas avec le passé simple est effectivement un indice qu'ils font partie d'une même catégorie, mais leur incompatibilité avec le subjonctif ne signifie pas forcément qu'ils font aussi partie de la même catégorie que lui. Il ne faut pas négliger qu'il soit possible que ces deux grammèmes soient des temps et que le subjonctif ne se combine tout simplement pas aux temps.

Nous croyons que les critères basés sur la combinatoire ne suffisent pas à classer les grammèmes en catégories flexionnelles, et qu'il faut tenir compte des sens qu'ils expriment. Or, le futur simple, même si on lui connaît des emplois modaux, sert la plupart du temps à situer un fait dans le futur, et on peut supposer que c'est ce sens qui vient spontanément aux locuteurs quand ils pensent à cette forme (il s'agit, dans nos termes, de son acception de base). Ce sens temporel n'a pas de parenté sémantique évidente avec celui (s'il existe) du subjonctif. Il n'y a donc pas lieu à notre avis de classer le futur simple dans la même catégorie que le subjonctif. Quant au passé simple, de l'aveu même de l'auteur, il ne sert jamais qu'à situer les faits dans le passé et ne peut pas porter de sens (purement) modal (on peut à la rigueur y reconnaître, comme le fait Touratier, une composante sémantique de 'factuel'). Il nous semble donc très bizarre de le classer parmi les modes plutôt que les temps.

6.2 Les critères basés sur le sens

Nous avons vu au début de ce chapitre (§ 1, p. 46) que Mel'čuk (1993: p. 261) définit le grammème comme un élément d'une catégorie flexionnelle. De façon plus générale, il définit les catégories comme des ensembles de significations qui s'excluent mutuellement dans une même position sémantique ou logique. C'est

donc principalement sur la base de critères sémantiques que Mel'čuk regroupe les grammèmes en catégories flexionnelles¹³. À partir de sa définition des catégories dans la langue, on peut formuler le critère suivant.

Critère de l'exclusion mutuelle sémantique

Les grammèmes d'une même catégorie flexionnelle expriment des sens mutuellement exclusifs.

Le problème avec une approche basée sur la sémantique est qu'elle n'est pas économique, puisqu'elle force la création aux niveaux syntaxiques et morphologiques (profonds et de surface) d'éléments distincts qui ont pourtant exactement le même comportement. Par exemple, si on se base essentiellement sur des critères sémantiques, il faut distinguer aux niveaux syntaxiques et morphologiques au moins deux grammèmes pour le conditionnel: un qui appartient au mode (pour le conditionnel de politesse [*J'en reprendrais bien un peu*]) et un autre qui appartient au temps (pour le «futur dans le passé» [*Il a dit qu'il jouerait avec nous ce soir*]). Le problème se complique encore quand on considère d'autres sens du conditionnel, par exemple la réserve [*Les voleurs se seraient enfuis à bord d'une fourgonnette bleue*]: faut-il y voir un grammème séparé du premier, ou une acception du même grammème? Pourtant, à partir du niveau syntaxique, jusqu'à la réalisation de la phrase en surface, tous ces grammèmes se comportent très exactement de la même façon. Il en résulte une duplication inutile des règles à tous les niveaux. Pour palier à ce problème, nous proposons le critère suivant.

Critère de la similitude sémantique

L'acception de base des grammèmes d'une même catégorie flexionnelle présente une parenté sémantique évidente.

En ne considérant que la grammie de base associée à un grammème, nous évitons le problème auquel s'expose Mel'čuk, ce qui nous permettra d'obtenir un modèle plus économique.

13. Il ne se base toutefois pas exclusivement sur de tels critères.

La méthodologie de Martinet, basée sur la combinatoire, et celle de Mel'čuk, basée surtout sur le sens, mènent à des descriptions linguistiques très différentes. Notre méthodologie est à cheval sur les deux : les grammèmes d'une même catégorie flexionnelle doivent être mutuellement exclusifs, avoir une combinatoire similaire, et leur acception de base doit avoir une parenté sémantique avec celle des autres grammèmes de cette catégorie.

7 Grammème et phraséologie

L'étude des grammèmes peut être compliquée par divers phénomènes de phraséologisation. On sait depuis longtemps que les expressions lexicales peuvent se phraséologiser ; par exemple, le sens de 'SE FAIRE UN SANG D'ENCRE' ne contient pas ceux de SANG ou de ENCRE. Cependant, les expressions non-compositionnelles ne sont pas forcément lexicales. De même que les lexies peuvent avoir un signifiant polylexémique, les grammies peuvent avoir un signifiant polygrammémique. Beck (2007) a démontré, dans le cadre de la TST, que le traitement de certains phénomènes morphologiques comme des cas de phraséologie permet une plus grande économie dans les modèles linguistiques. Il donne notamment l'exemple du mode irréel en totonac (une langue amérindienne parlée dans l'état de Puebla au Mexique), qui n'a pas de morphème propre, mais qui s'exprime par la combinaison des morphèmes de temps passé et d'aspect perfectif. Pour Beck, le problème se traite dans la transition entre les niveaux morphologiques profond et de surface, c'est-à-dire qu'au niveau syntaxique, on retrouve le grammème **irréel**. Dans le cadre de notre travail cependant, nous ne considérons que l'interface sémantique-syntaxe. Nous traiterons donc ce type de phrasèmes dans cette interface, c'est-à-dire qu'au niveau syntaxique de surface, nous aurions pour le totonac les grammèmes **passé** et **perfectif**, mais pas ***irréel**. Pour cette raison, nous parlerons plutôt de phrasèmes grammémiques, et non morphologiques. Un **phrasème grammémique** est un signe représentable dans son signifiant en termes de plus d'un grammème, mais dont le signifié n'est pas compositionnel. Par exemple, en français, nous verrons que le signifiant du conditionnel à valeur modale (de politesse ou de réserve)

peut être décrit par la combinaison de deux grammèmes, correspondant respectivement aux morphèmes –R– et –AI–, alors que son signifié n'est pas la composition du sens de ces deux grammèmes.

Il faut remarquer que la phraséologie n'implique pas forcément des éléments de même nature. Par exemple, un signe peut avoir comme signifiant une combinaison de grammèmes et de lexèmes, comme dans la locution «CESSEZ-LE-FEU», où le grammème **impératif** fait partie intégrante du signifiant, conjointement avec les lexèmes CESSER, LE et FEU.

Les phrasèmes grammémiques peuvent facilement brouiller les pistes dans l'étude du système flexionnel de la langue, et il faut toujours s'assurer, quand on cherche à décrire un grammème, qu'on a bien affaire à un grammème simple, et non à un phrasème grammémique ou à une partie d'un phrasème. Ainsi, par exemple, les emplois modaux du conditionnel en français ne doivent pas être pris en compte pour la classification des grammèmes qui en composent le signifiant (et qui correspondent aux morphèmes –R– et –AI–).

8 Synthèse

En résumé, nous avons distingué les grammèmes profonds, qui appartiennent à la représentation syntaxique profonde, et les grammèmes de surface, qui relèvent du niveau syntaxique de surface. Les grammèmes ne sont pas des signes. Il sont du niveau d'abstraction du vocable et il faut en distinguer les différentes acceptions, les grammies (qui, elles, sont des signes). Parmi les grammies associées à un grammème, une est considérée comme la grammie de base. On peut l'identifier grâce à une variété de critères basés sur le sens et sur la combinatoire des signes en jeu.

C'est la grammie de base des grammèmes qui doit guider la construction du modèle linguistique. C'est en effet en fonction du sens et de la combinatoire de cette dernière que les grammèmes sont regroupés en catégories flexionnelles. Les grammies de base des grammèmes d'une même catégorie flexionnelle doivent être

mutuellement exclusives et avoir une combinatoire similaire, en plus de présenter une parenté sémantique évidente.

Enfin, il faut considérer que les grammèmes peuvent se phraséologiser, ce qui peut brouiller les pistes. Il faut donc toujours s'assurer qu'on a affaire à un grammème simple, et non à une composante d'un phrasème grammémique.

V. Revue critique de quelques modèles de la conjugaison française

Le littérature touchant les questions de temps et d'aspect verbaux dans les langues est énorme. Le simple fait d'en dresser l'inventaire constitue en soi une tâche colossale, tâche à laquelle Binnick (2001) s'est attelé pour produire une bibliographie beaucoup plus volumineuse que cette thèse, contenant plus de 6 600 publications principalement en anglais, en français et en allemand. Il n'a pourtant inclus que les ouvrages publiés après 1947, sauf quelques exceptions, et exclu la plupart de ceux qui se situaient dans le cadre de la sociolinguistique, de la dialectologie, de la linguistique informatique, de la structure narrative, etc. Il a recensé environ 300 publications par année tout au long des années 90, mais beaucoup moins dans les trois ou quatre années précédant la publication de sa bibliographie. On peut donc supposer qu'il manque de nombreuses références entre 1998 et 2001, sans compter tout ce qui s'est publié depuis 2001 !

Une telle masse d'information ne peut pas être assimilée en quelques années. Il va de soi que nous n'avons pas tout lu, ni même toutes les publications importantes. Nous avons presque exclusivement orienté nos recherches bibliographiques vers ce qui traitait du temps (et de l'aspect) en français ou dans les langues en général. Même en nous restreignant à ces ouvrages, il est impossible d'offrir ici un panorama exhaustif de ce qui s'est écrit sur le temps verbal en français. Nous nous contenterons de présenter brièvement quelques-uns des modèles que nous avons rencontrés. Il ne faut pas croire que nous avons forcément choisi les plus marquants. Nous laissons de côté plusieurs auteurs qui ont ou ont eu une influence notable. Rien, par exemple, sur Reichenbach, Partee ou encore Comrie, qui ont eu un impact considérable dans le domaine de l'étude du temps grammatical dans la langue en général. Rien non plus sur Verkuyl, Vet, Molendijk ou encore Veters, qui sont pourtant des contemporains de gros calibre dans le domaine du temps grammatical en français. Nous avons choisi de comparer uniquement des ouvrages qui offrent un modèle complet de la flexion verbale plutôt que de se pencher sur un

sous-ensemble des formes verbales. Nous aurons néanmoins l'occasion de revenir sur d'autres auteurs au cours de notre discussion des différentes catégories flexionnelles de notre propre modèle, là où cela sera pertinent. Par ailleurs, le lecteur intéressé pourra consulter notamment Vet (1980) et Veters (1996), qui passent en revue des auteurs que nous ne présentons pas dans cette thèse.

Rappelons que ce qui nous intéresse dans le cadre de cette thèse est de construire un modèle basé sur des catégories flexionnelles composées de grammèmes répondant aux caractéristiques définies au Chapitre IV. C'est dans cette optique que nous présentons les travaux de nos prédécesseurs. Cependant, tous les auteurs présentés ici ne raisonnent pas forcément en termes de catégories flexionnelles et de grammèmes. Les cadres théoriques variant énormément, les questionnements ne sont pas les mêmes pour chacun. Il est assez fréquent notamment qu'on propose des décompositions sémantiques des formes verbales plutôt qu'un système flexionnel au sens où nous l'entendons. Afin de pouvoir comparer et critiquer selon notre perspective particulière les différents modèles discutés ici, nous avons choisi de les représenter en termes de grammèmes et de catégories flexionnelles, bien que cela introduise inévitablement des distorsions. Évidemment, la distinction entre grammème profond et grammème superficiel (*cf.* Ch. IV, § 1, p. 46) n'est pas vraiment pertinente ici puisqu'elle est propre à notre cadre de travail et qu'aucun des auteurs présentés ne s'en préoccupe.

Partout nous faisons abstraction de la voix et des catégories d'accord (personne, nombre et genre). Nous nous contentons d'indiquer à quels grammèmes se combinent la personne et le nombre verbaux, sans nous soucier de savoir s'il s'agit d'une seule catégorie flexionnelle ou de deux catégories différentes — ce qui varie d'un auteur à l'autre. L'accord des participes est complètement laissé de côté ici.

Avant de commencer, rappelons qu'un grammème ne s'exprime pas forcément par un morphème séparé; il ne faut donc pas comprendre que nous cherchons à opérer une division morphématique des formes verbales. Il n'est peut-être pas inutile de rappeler également quelques conventions d'écriture qui prennent tout leur

sens dans ce chapitre. Nous utilisons les termes traditionnels, sans police de caractères particulière, pour faire référence à une forme verbale sans égard à son analyse morphologique: «imparfait» renvoie aux *formes mangeait, mangions*, etc. Nous utilisons une police particulière lorsque nous faisons référence à un grammème: «**imparfait**» est le nom d'un *grammème* (dont le signifié pourra varier d'un modèle à l'autre). Enfin, nous utilisons encore une autre police pour les morphèmes: «-AI-» est le nom d'un *suffixe* (qui comprend les allomorphes *-i-*, aux première et deuxième personnes du pluriel [*mangions, mangiez*], et *-ai-*, aux autres personnes [*mangeais, mangeait, mangeaient*]). Il ne faut surtout pas confondre les trois, et il n'y a pas forcément de relation entre eux. Dans certains modèles, par exemple, le conditionnel est formé d'un suffixe *-R-* et d'un suffixe *-AI-*, sans que ce *-AI-* ne corresponde à un grammème nommé **imparfait**. Dans d'autres modèles, le grammème **imparfait** peut s'exprimer par un morphème autre que *-AI-*. Et chez d'autres auteurs encore, les formes de l'imparfait n'expriment pas un grammème appelé **imparfait**, mais autre chose!

1 Le modèle traditionnel

Il y a un large consensus parmi les grammaires scolaires et les ouvrages de référence les plus usuels (*Le bon usage, Le Bescherelle, Le Larousse des verbes, Le Petit Robert*, etc). Les tableaux de conjugaison y sont pratiquement identiques et proposent un modèle relativement plat qui ne cherche pas à identifier des grammèmes et des catégories flexionnelles. Le but de ces ouvrages est plutôt de classer et de nommer les formes verbales, puis d'en décrire les différents usages (sémantiques et pragmatiques). Pour représenter le modèle traditionnel sous forme de catégories flexionnelles et de grammèmes, on peut supposer dans un premier temps que les modes et les temps correspondent chacun à un grammème (même si, nous le verrons plus loin, les commentaires qu'on trouve dans ces grammaires de référence laissent voir un système un peu mieux articulé). On obtient alors un système où interviennent peu de catégories flexionnelles.

Dans les tableaux de conjugaison française traditionnels, mis à part les catégories d'accord (et la voix, qui ne fait habituellement pas partie de ces tableaux), on ne peut distinguer en apparence que deux catégories flexionnelles: le mode et le temps.

Les modes sont divisés en modes personnels (**indicatif**, **conditionnel**, **subjonctif** et **impératif**) et modes impersonnels (**infinitif**, **participe** et **gérondif**) selon qu'ils portent ou non la flexion en personne et en nombre. Tous se combinent à des grammèmes de temps, mais en fait, seul le mode **indicatif** peut se combiner à tous les temps; les autres modes ne se combinent qu'aux grammèmes **présent** et **passé (composé)**, sauf le subjonctif qui se combine également aux grammèmes **imparfait** et **plus-que-parfait**.

La catégorie de temps contient un grammème différent pour chaque groupe de formes de l'indicatif: **présent**, **passé simple**, **imparfait**, **futur simple**, **passé (composé)**, **passé antérieur**, **plus-que-parfait** et **futur antérieur**. Leurs sens se décrivent en termes de relations temporelles et d'aspect. Par exemple, le passé simple, l'imparfait et le passé composé situent tous les faits dans le passé mais se distinguent par leurs valeurs aspectuelles (nous n'insisterons pas sur ce point ici; voir le modèle de Imbs ci-dessous ou encore Comrie 1976: pp. 126-127).

Notons que dans plusieurs ouvrages, surtout les plus récents (par exemple, Grevisse 1993, ou encore Riegel, Pellat & Rioul 1994), le conditionnel est considéré comme un temps de l'indicatif plutôt que comme un mode. Il y a alors dans ces modèles deux grammèmes de temps de plus: **conditionnel (présent)** et **conditionnel passé**.

Le futur périphrastique est traité de façon ambiguë dans la tradition. Il est presque systématiquement absent des tableaux de conjugaison des grammaires scolaires et des ouvrages de référence, mais pourtant presque tous en reconnaissent l'existence et en disent quelques mots. On l'a d'abord appelé «futur proche», mais il est maintenant généralement admis qu'il ne situe pas forcément les faits dans un futur rapproché, mais plutôt qu'il «marque un futur vu du présent, souvent un futur

proche, parfois un futur plus lointain mais considéré comme inéluctable» (Grevisse 1993: p. 789). Étant donné son traitement timide dans les ouvrages traditionnels, nous l'excluons de notre discussion ici (nous aurons l'occasion d'y revenir plus loin, Ch. VI, § 5.4, p. 170).

La Figure 11 et le Tableau VI présentent les catégories flexionnelles qu'on peut déduire du modèle traditionnel, leurs grammèmes et leur combinatoire, ainsi qu'un exemple pour chaque combinaison.

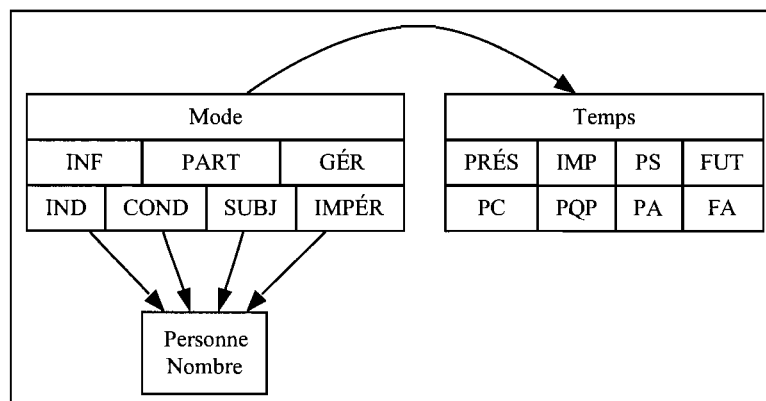


Figure 11 — Les catégories flexionnelles dans le modèle traditionnel

Mode	Temps	Forme
indicatif	présent	<i>sait</i>
	imparfait	<i>savait</i>
	passé simple	<i>sut</i>
	futur simple	<i>saura</i>
	passé composé	<i>a su</i>
	plus-que-parfait	<i>avait su</i>
	passé antérieur	<i>eut su</i>
	futur antérieur	<i>aura su</i>
conditionnel	présent	<i>saurait</i>
	passé	<i>aurait su</i>

Tableau VI — Le classement traditionnel des formes

Mode	Temps	Forme
subjonctif	présent	<i>sache</i>
	passé	<i>ait su</i>
	imparfait	<i>sût</i>
	plus-que-parfait	<i>eût su</i>
impératif	présent	<i>sache !</i>
	passé	<i>aie su !</i>
infinitif	présent	<i>savoir</i>
	passé	<i>avoir su</i>
participe	présent	<i>sachant</i>
	passé	<i>su / ayant su</i>
gérondif	présent	<i>en sachant</i>
	passé	<i>en ayant su</i>

Tableau VI — Le classement traditionnel des formes

Ce modèle est en un sens économique, puisqu'il ne fait appel qu'à deux catégories flexionnelles, chacune ne comportant pas plus de huit grammèmes. Cependant, nous croyons que ces catégories peuvent être raffinées.

Commençons par le mode. Il nous semble que l'opposition entre modes personnels (**indicatif**, **conditionnel**, **subjonctif** et **impératif**) et impersonnels (**infinitif**, **participe** et **gérondif**) est plus profonde que le simple fait que les uns expriment la personne et les autres non. Il y a des différences sémantiques importantes. D'ailleurs, on sent bien le malaise des grammairiens dans les passages suivants, quand ils tentent de définir ce qu'est la catégorie de mode :

L'identification des modes aux modalités ne permet pas de caractériser chacun d'eux par des propriétés vraiment distinctives. En premier lieu, on ne voit pas quelles modalités pourraient être exprimées par l'infinitif et par le participe.

Riegel, Pellat & Rioul (1994: p. 287)

[...] il est difficile de donner une définition précise de la notion de mode. [...] Les modes impersonnels sont, entre eux, plus homogènes. Ils permettent en effet de conférer à un verbe (muni éventuellement de tous ses compléments et,

V. Revue critique de quelques modèles de la conjugaison française

parfois, de son sujet) les fonctions généralement exercées par un mot d'une autre classe: nom pour l'infinitif, adjectif pour le participe, adverbe pour le gérondif.

Bescherelle (1998: § 156)

Visiblement, les modes impersonnels ne cadrent pas bien et empêchent de définir clairement ce qu'est un mode. En effet, leur fonction est clairement différente de celle des modes personnels. Alors que les modes en général sont supposés refléter l'attitude du locuteur par rapport à l'énoncé (portant ainsi un sens d'ordre pragmatique), seuls les modes personnels le font véritablement. Les modes impersonnels, eux, ne portent aucun sens. Ils ont une fonction plus abstraite n'ayant à voir qu'avec l'organisation syntaxique de la phrase. Ces deux types de grammèmes forment donc deux classes sémantiquement bien distinctes. Qui plus est, ils s'opposent également par leur combinatoire: les modes personnels acceptent la flexion en personne alors que les modes impersonnels, non. Pour ces deux raisons, nous croyons qu'il faut voir ici deux catégories flexionnelles plutôt qu'une seule. Nous y reviendrons plus loin (Ch. VI, § 1, p. 116 et Ch. VI, § 3, p. 134).

Pour ce qui est du temps, notons qu'à chacun des grammèmes qui s'expriment de façon synthétique (**présent, imparfait, passé simple, futur simple**) correspond un temps dont le signifiant comprend l'auxiliaire AVOIR <ÊTRE> et dont le sens par rapport à la forme non composée est toujours sensiblement le même¹, de sorte qu'on peut formuler l'équivalence $\frac{\text{PRES}}{\text{PC}} = \frac{\text{IMP}}{\text{PQP}} = \frac{\text{PS}}{\text{PA}} = \frac{\text{FUT}}{\text{FA}}$. Nous voyons là matière à postuler une catégorie flexionnelle supplémentaire dont les grammèmes se combineraient aux grammèmes de temps pour former les temps non composés d'une part et composés d'autre part (et même, pourquoi pas, surcomposés). D'ailleurs, on retrouve la même opposition sémantique et formelle entre les formes dites «présent» et «passé» des autres modes (par exemple, l'infinitif passé est à l'infinitif ce que le futur antérieur est au futur simple). Or, il nous semble que cette opposition n'est pas vraiment d'ordre temporel, ou du moins, que les temps

1. Ce n'est pas tout à fait exact puisque le passé composé et le plus-que-parfait peuvent apporter un sens par rapport au présent et à l'imparfait que le passé antérieur et le futur antérieur n'ont pas par rapport au passé simple et au futur simple. Nous y reviendrons plus loin (Ch. VI, § 5, p. 164).

«présent» et «passé» de ces modes n'ont pas le même sens qu'à l'indicatif². Par exemple, le «subjonctif présent» de *Il ne fallait pas qu'il fasse un seul mouvement* ne situe clairement pas les faits dans le présent. Il faut le comprendre comme indiquant plutôt la concomitance. Puisque, en plus de cette différence sémantique, ce «présent» ne commute pas avec les mêmes grammèmes qu'à l'indicatif, il ne s'agit pas du même grammème **présent** à l'indicatif et au subjonctif.

Ces faits ont bien entendu été notés très tôt (déjà, la grammaire de Port-Royal en faisait mention) et la plupart des grammaires plus ou moins récentes semblent y voir comme nous une catégorie flexionnelle, bien que cela ne se reflète pas dans les tableaux de conjugaison. Toutefois, il est généralement accepté au moins depuis Benveniste (1959) que l'auxiliaire AVOIR <ÊTRE> peut avoir deux valeurs distinctes: celle d'accomplissement et celle d'antériorité. On hésite néanmoins dans la tradition à voir là deux grammèmes bien distincts. Or, nous croyons que la clé des temps surcomposés et de certains autres phénomènes réside précisément dans la reconnaissance de deux grammèmes profonds correspondant à chacune de ces valeurs et appartenant à deux catégories séparées. Nous y reviendrons au chapitre suivant (Ch. VI, § 4, p. 144).

Autre élément qui n'apparaît pas dans les tableaux de conjugaison mais qui est pourtant un thème récurrent dans les grammaires: la distinction entre temps absolus et temps relatifs. On trouve chez Veters (1996) un historique critique des différentes positions sur le sujet, de la grammaire de Port-Royal aux auteurs les plus récents. Nous nous contenterons ici de quelques points essentiels.

On classe traditionnellement les temps qui situent les faits directement par rapport au moment d'énonciation dans la catégorie des temps absolus, alors que les temps relatifs sont ceux qui situent les faits par rapport à un point de référence qui est antérieur ou postérieur au moment d'énonciation (il s'agit souvent d'un autre fait qu'on retrouve dans le discours). L'inventaire de chacune de ces deux classes

2. Nous ne savons pas vraiment comment interpréter ce «passé» en termes de grammèmes. Est-ce le même grammème que pour le passé composé ou le passé simple, ou est-ce un troisième grammème? Ou encore, est-ce un terme qui recouvre plusieurs grammèmes?

varie légèrement d'un auteur à l'autre. Comparons les classements des formes de l'indicatif de Girard (1747) et de Destutt de Tracy (1803):

	Absolus	Relatifs
Girard	<i>sait</i> <i>a su</i> <i>sut</i> <i>saura</i>	<i>savait</i> <i>avait su</i> <i>eut su</i> <i>aura su</i>
Destutt de Tracy	<i>sait</i> <i>a su</i> <i>a eu su</i> <i>saura</i> <i>aura su</i> <i>aura eu su</i>	<i>savait</i> <i>avait su</i> <i>avait eu su</i> <i>saurait</i> <i>aurait su</i> <i>aurait eu su</i>

Tableau VII — Les temps absolus et relatifs dans la tradition

Quelle est la nature exacte de ce genre de dichotomie? Faut-il y voir deux catégories flexionnelles de temps qui sont en concurrence, ou est-ce un simple regroupement de grammèmes, au sein d'une unique catégorie de temps, selon des similarités sémantiques? Ou encore, faut-il considérer que «absolu» et «relatif» sont les deux grammèmes d'une catégorie flexionnelle supplémentaire qui se combinerait à celle du temps?

S'il s'agissait de deux catégories flexionnelles de temps concurrentes, de sorte que le modèle flexionnel contienne en fait trois catégories flexionnelles (mode, temps absolu et temps relatif), alors les catégories de temps absolu et de temps relatif seraient mutuellement exclusives, ce qui compliquerait la modélisation du caractère obligatoire de la flexion. En effet, si on a deux catégories de temps concurrentes, il faut alors faire appel à des critères syntaxiques, et non plus seulement morphologiques, pour déterminer quelles sont les catégories flexionnelles obligatoires. Il ne suffit plus de dire qu'un verbe à l'indicatif doit être fléchi en temps (ce qui peut se modéliser simplement au niveau morphologique), mais il faut encore tenir compte du contexte syntaxique; on dira par exemple qu'un verbe à l'indicatif dans une proposition principale doit être fléchi en temps absolu, alors que dans une proposition subordonnée il doit être fléchi en temps relatif³. Pour cette raison, il ne

nous semble pas souhaitable que le locuteur ait le choix entre deux catégories flexionnelles. Ce classement ne doit donc pas correspondre à deux catégories flexionnelles concurrentes.

L'hétérogénéité des formes dans le classement de Girard⁴ laisse croire qu'il s'agit simplement d'un regroupement de grammèmes sans conséquence sur la structure du système flexionnel: il s'agit d'un commentaire sur le signifié des grammèmes de la catégorie de temps. Par contre, le modèle de Destutt de Tracy est intéressant en ce que la différence sémantique entre les temps absolus et les temps relatifs va de pair avec une différence formelle régulière (on ajoute le suffixe –AI– à un temps absolu pour obtenir le temps relatif correspondant). On pourrait donc y voir une catégorie flexionnelle qui contiendrait deux grammèmes, **absolu** (au signifiant nul) et **relatif** (dont le signifiant serait le suffixe –AI–), qui se combineraient aux grammèmes de temps (cette catégorie étant ici réduite aux grammèmes **présent, passé composé, passé surcomposé, futur, futur antérieur et futur antérieur surcomposé**). N'ayant malheureusement pas accès au texte original⁵, il nous est impossible de dire si c'est réellement ce que l'auteur avait en tête. Cependant, on peut constater que ce n'est au moins pas ce qu'on en a retenu puisque les auteurs que nous connaissons qui ont proposé des classements similaires ne semblaient pas voir là une catégorie flexionnelle. Il est possible que le fait que le passé simple et le passé antérieur brisent cette belle symétrie, comme le note Vettters (1996), en ait découragé plusieurs de stipuler un grammème **relatif**. Or, nous proposons justement que cette dichotomie entre temps absolus et temps relatifs est due à l'existence d'une catégorie flexionnelle, que nous appelons «décalage», qui se combine aux temps, et dont le grammème **décalé** s'exprime justement par le suffixe –AI–. Nous y reviendrons au prochain chapitre (Ch. VI, § 6, p. 177) et nous

-
3. La règle, en fait, doit être beaucoup plus complexe puisque les temps relatifs peuvent apparaître dans une principale [*Tiens, je te pensais à Québec, toi*] et les temps absolus peuvent aussi se trouver dans une subordonnée [*Je te jure que j'ai rien touché*].
 4. La différence formelle entre un temps relatif et son correspondant absolu est parfois l'imparfait, parfois l'auxiliaire AVOIR, auxiliaire qui se trouve d'ailleurs aussi dans le signifiant de certains temps absolus.
 5. Nous connaissons le classement de Destutt de Tracy par Vettters (1996), qui lui-même le tenait d'un article d'Yvon (1951). Le lecteur nous pardonnera ce brocantage.

V. Revue critique de quelques modèles de la conjugaison française

verrons que le passé simple et le passé antérieur, bien qu'ils brisent effectivement la symétrie des formes, ne posent pas vraiment problème.

La distinction entre temps absolus et relatifs a perduré et on l'a souvent reprise⁶. On la retrouve également dans d'autres langues, si bien qu'aujourd'hui il s'agit d'une distinction largement acceptée en linguistique générale (*cf.* Comrie 1985). Parmi les auteurs qui l'ont repris, on trouve Imbs, qui a construit son modèle essentiellement sur cette base.

2 Imbs

Imbs (1960) propose un modèle où le temps est représenté non pas de façon unidimensionnelle, mais en faisant appel à trois axes. Un premier axe est ancré dans le présent et divise le temps entre passé, présent et futur. Le passé et le futur sont à leur tour eux-même traversés par des axes secondaires qui fournissent des temps antérieurs ou postérieurs au passé et au futur.

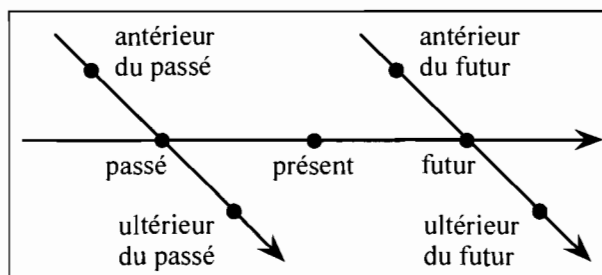


Figure 12 — Les trois axes temporels de Imbs

L'auteur ne va toutefois pas jusqu'à postuler deux (ou trois) catégories flexionnelles de temps différentes. On a ici un système à sept temps: **passé** (*a mangé, mangeait et mangea*), **présent** (*mange*), **futur** (*mangera*), **antérieur du passé** (*a eu mangé, avait mangé et eut mangé*), **ultérieur du passé** (*mangerait*), **antérieur du futur** (*aura mangé*) et **ultérieur du futur** (*mangera*).

6. On l'a éventuellement reformulée en termes de temps « déictiques » vs « anaphoriques » (voir notamment Partee 1973, 1984).

L'imparfait, le passé composé et le passé simple (ainsi que le plus-que-parfait, le passé surcomposé et le passé antérieur) se distinguent par l'aspect. L'imparfait exprime le grammème d'aspect **duratif-itératif**, le passé composé, le grammème **accompli**, et le passé simple, le grammème **ponctuel**. Imbs voit également une opposition aspectuelle entre les formes simples et composées aux autres temps et modes. Il a donc un grammème **inaccompli** pour les formes simples du présent, de l'infinitif, etc., s'opposant aux formes composées correspondantes, qui expriment le grammème **accompli**. Cette opposition peut, selon le contexte, s'interpréter comme une opposition entre simultanéité et antériorité: ces deux grammèmes ont un sens vague.

Il n'est pas clair s'il faut voir dans ce modèle une ou deux catégories d'aspect. L'opposition **duratif-itératif / accompli / ponctuel** n'est utile que pour les temps **passé** et **antérieur du passé**⁷. Partout ailleurs, on ne trouve que l'opposition **accompli / inaccompli**. Puisqu'il serait bizarre de placer le grammème **accompli** dans deux catégories différentes, nous représentons le modèle de Imbs avec une seule catégorie d'aspect.

Les modes de Imbs sont ceux de toujours, sauf que le conditionnel est ici vu comme un temps de l'indicatif et non comme un mode. Il y a donc dans son modèle les grammèmes de mode **indicatif**, **subjonctif**, **impératif**, **infinitif**, **gérondif** et **participe**. Nous laissons de côté ici son traitement des temps du subjonctif dans le style soutenu pour ne considérer que les formes du subjonctif présent (*sache*) et du subjonctif passé (*ait su*), qui s'opposent dans son modèle par l'aspect.

La Figure 13 et le Tableau VIII ci-dessous présentent de façon schématique le modèle de Imbs et donnent, pour chaque combinaison de grammème possible, un exemple.

7. Imbs fait remarquer que le sens de ces grammèmes aspectuels se retrouve également au présent. Par contre, ils n'entraînent aucun changement dans la forme, aussi croyons-nous que cette opposition n'est pas pertinente au présent.

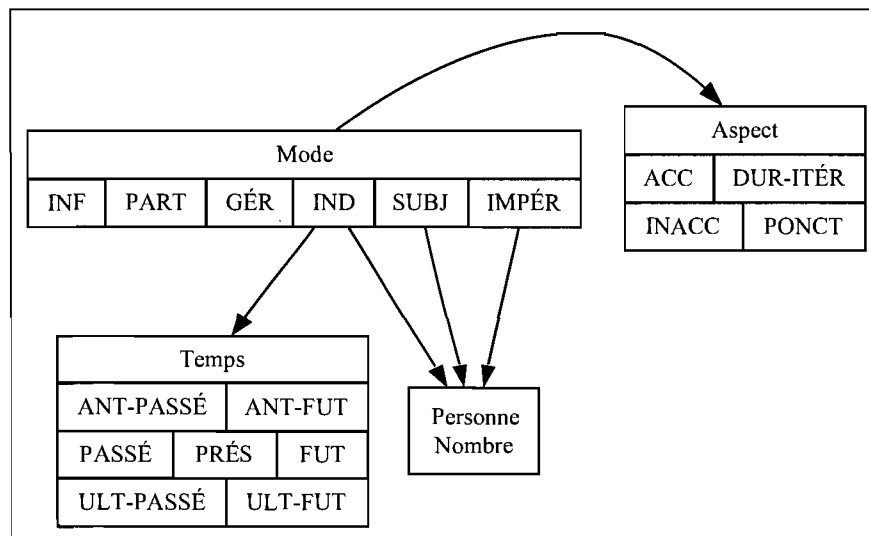


Figure 13 — Les catégories flexionnelles dans le modèle de Imbs

Mode	Temps	Aspect	Forme
indicatif	présent	inaccompli	<i>sait</i>
		accompli	<i>a su</i>
	passé	accompli	<i>a su</i>
		duratif-itératif	<i>savait</i>
		punctuel	<i>sut</i>
	antérieur du passé	accompli	<i>a eu su</i>
		duratif-itératif	<i>avait su</i>
		punctuel	<i>eut su</i>
	ultérieur du passé	inaccompli	<i>saurait</i>
		accompli	<i>aurait su</i>
	futur	inaccompli	<i>saura</i>
		accompli	<i>aura su</i>
	antérieur du futur	inaccompli	<i>aura su</i>
		accompli	<i>aura eu su</i>
	ultérieur du futur	inaccompli	<i>saura</i>
		accompli	<i>aura eu su</i>

Tableau VIII — Le classement des formes selon Imbs

V. Revue critique de quelques modèles de la conjugaison française

Mode	Temps	Aspect	Forme
subjonctif		inaccompli	<i>sache</i>
		accompli	<i>ait su</i>
impératif		inaccompli	<i>sache!</i>
		accompli	<i>aie su!</i>
infinitif		inaccompli	<i>savoir</i>
		accompli	<i>avoir su</i>
participe		inaccompli	<i>sachant</i>
		accompli	<i>su / ayant su</i>
gérondif		inaccompli	<i>(en) sachant</i>
		accompli	<i>(en) ayant su</i>

Tableau VIII — Le classement des formes selon Imbs

Certaines formes surcomposées reconnues comme grammaticales par Imbs ne trouvent pas leur place dans ce modèle. Par exemple, comment faut-il analyser la forme *aurait eu chanté*? Le conditionnel du premier auxiliaire indique un grammème **ultérieur du passé**, mais que faire du double auxiliaire? Imbs y voit, à juste titre sans doute, une marque à la fois d'antériorité et d'accomplissement. Mais quels sont les grammèmes en jeu? S'agit-il deux fois du même grammème **accompli**, une fois avec une interprétation d'antérieur et une fois avec une interprétation d'accompli? Nous trouvons douteuse l'idée qu'un grammème se combine avec lui-même. Et encore, comment expliquer qu'une seule des trois formes du temps **antérieur du passé**, à savoir le plus-que-parfait, puisse être composée une fois de plus (*avait chanté* → *avait eu chanté*, mais *eut chanté* → **eut eu chanté* et *a eu chanté* → **a eu eu chanté*)? D'ailleurs, comment doit-on représenter ce plus-que-parfait surcomposé? Il doit s'agir en principe d'un accompli par rapport au plus-que-parfait, mais puisque le plus-que-parfait exprime déjà lui-même le grammème **duratif-itératif**, qui fait partie de la même catégorie flexionnelle que **accompli**, il y a un problème.

Le manque d'homogénéité que présente la catégorie d'aspect chez Imbs (les oppositions ne sont pas les mêmes pour les temps **passé** et **antérieur du passé** que pour les autres temps et mode, comme nous l'avons mentionné plus haut) nous semble gênant. Nous ne sommes pas convaincu qu'il s'agisse vraiment d'une seule catégorie. Nous proposons plus loin (Ch. VI, § 6, p. 177 et Ch. VI, § 6.1.2, p. 181) une alternative à la catégorie d'aspect pour expliquer les différences entre les formes de l'imparfait, du passé composé et du passé simple (ainsi que les formes composées et surcomposée correspondantes).

Ce que nous retenons principalement du modèle de Imbs est qu'il postule plus d'un axe temporel. Nous défendons plus loin (Ch. VI, § 6, p. 177) un système similaire où interviennent deux axes, à chacun desquels correspond une catégorie flexionnelle distincte.

3 Les fonctionnalistes

3.1 Martinet

Martinet consacre une cinquantaine de pages de sa *Grammaire fonctionnelle du français* (1979) aux verbes et à leur conjugaison. Son modèle s'appuie sur une approche générale de la langue rigoureuse et systématique (exposée notamment dans Martinet 1970, 1979), ce qui en fait un modèle particulièrement intéressant à étudier. En outre, il fait clairement appel à des grammèmes regroupés en catégories flexionnelles⁸. Sa théorie de la flexion repose essentiellement sur deux principes fondamentaux : les grammèmes d'une même catégorie doivent **1)** être mutuellement exclusifs et **2)** avoir une combinatoire semblable. C'est donc presque exclusivement en fonction de leur combinatoire que les grammèmes sont regroupés en catégories flexionnelles.

En excluant la voix et les catégories d'accord, on compte dans ce modèle quatre catégories flexionnelles : le mode, le temps, l'aspect et la vision.

8. La terminologie de Martinet diffère de la nôtre : il parle plutôt de monèmes et de classes de monèmes (ou modalités).

La catégorie de mode chez Martinet comprend quatre grammèmes: **infinitif**, **participe**, **impératif** et **subjonctif**. Le conditionnel n'est pas considéré ici comme un mode, mais comme la combinaison d'un grammème de temps et d'un de vision (nous y reviendrons). Le participe présent et le participe passé ne correspondent pas à deux grammèmes de mode différents, mais expriment tous les deux le même mode **participe**, le participe passé exprimant en plus le grammème d'aspect **parfait** (que nous verrons plus loin).

Là où Martinet se distingue particulièrement de la tradition, c'est dans son traitement de l'indicatif, qu'il ne considère pas comme un mode. En fait, il y voit carrément une absence de signe:

Ce que la tradition désigne comme «mode indicatif» n'a ni marque formelle, ni valeur distincte de celle du verbe nu. En conséquence, il ne constitue pas une unité linguistique particulière.

Martinet (1979: p. 111)

De façon similaire, le présent n'est pas considéré comme un temps, mais encore comme l'absence de signification temporelle. Ainsi la forme *chante* est-elle conçue comme un radical nu, sans aucun autre signe⁹. La catégorie de temps ne comprend donc pas de grammème correspondant au présent. Il n'y a que cinq temps: **passé** (qui correspond au suffixe –A– de l'imparfait), **prétérit** (pour le passé simple), **futur** (qui correspond au suffixe –R– du futur simple), **récent** (*il vient de chanter*) et **prochain** (*il va chanter*).

Conséquence de l'absence d'indicatif et de présent, les modes et les temps sont mutuellement exclusifs, à la seule exception du mode **subjonctif**, qui ne peut se combiner qu'au temps **passé** (ce qui donne le subjonctif imparfait).

Des cinq grammèmes de temps, seuls les trois premiers (**passé**, **prétérit** et **futur**), ainsi que tous les grammèmes de mode, peuvent se combiner à l'unique grammème d'aspect, **parfait**, qui s'exprime par l'auxiliaire AVOIR (ou ÊTRE). On

9. Sauf peut-être des signes prosodiques, que nous ne prenons pas en compte ici.

aura deviné que l'absence d'auxiliaire ne correspond pas selon Martinet à un deuxième grammème d'aspect qui s'opposerait à **parfait**.

Enfin, et c'est là encore une innovation par rapport à la tradition, Martinet postule l'existence d'une catégorie de vision, qui ne contient qu'un seul grammème: **décalé**. Il est associé aux mêmes signifiants que le temps **passé**:

Le temps passé et la vision décalée ne se distinguent en rien par la forme, et leurs valeurs sont souvent parallèles. Ils ne s'opposent que parce que le passé alterne avec les autres temps, tandis que le décalé coexiste avec eux.

Martinet (1979: p. 103)

La Figure 14 et le Tableau IX ci-dessous récapitulent les catégories flexionnelles et la combinatoire de leurs grammèmes d'après Martinet, et fournissent pour chaque combinaison possible un exemple. Précisons que Martinet ne dit pas explicitement à quels grammèmes peuvent se combiner les catégories d'accord (c'est pourquoi nous les indiquons en pointillés: il s'agit d'une extrapolation de notre part). Aussi, nous n'avons pas trouvé dans le texte d'indication claire sur le traitement de la forme simple de l'impératif. Nous l'indiquons ici comme portant le grammème **impératif**, bien que dans le tableau de Martinet (1979: pp. 100-101) cette forme ne soit associée à aucun grammème. L'auteur souligne le fait que l'impératif n'a pas de marque directe, si ce n'est l'absence de sujet, mais il lui reconnaît une valeur propre. Il doit donc s'agir d'un signe, et c'est pourquoi nous faisons cette petite entorse à son modèle, ce qui n'a de toute façon aucune incidence sur notre propos.

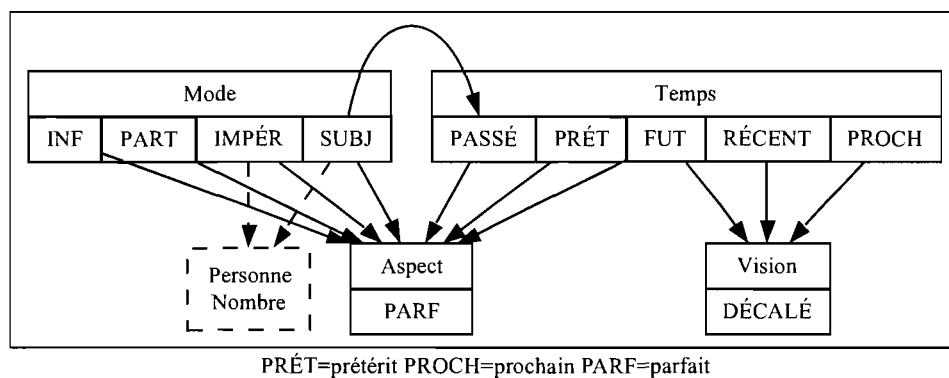


Figure 14 — Les catégories flexionnelles dans le modèle de Martinet

V. Revue critique de quelques modèles de la conjugaison française

Mode	Temps	Aspect	Vision	Forme
infinitif				<i>savoir</i>
		parfait		<i>avoir su</i>
participe				<i>sachant</i>
		parfait		ayant su
impératif				<i>sache!</i>
		parfait		aie su!
subjonctif				sache
		parfait		ait su
	passé			<i>sût</i>
		parfait		<i>eût su</i>
				sait
		parfait		a su
	passé			savait
		parfait		avait su
	prétérit			<i>sut</i>
		parfait		<i>eut su</i>
	futur			saura
			décalé	saurait
		parfait		aura su
			décalé	aurait su
	récent			vient de savoir
			décalé	venait de savoir
	prochain			va savoir
			décalé	allait savoir

Tableau IX— Le classement des formes selon Martinet

Une chose qui nous frappe chez Martinet est l'absence des formes surcomposées: son modèle ne permet pas d'en rendre compte. D'autres formes manquent également à l'appel: *il viendrait de chanter* (peut-être peu usitée, mais

quand même grammaticale) ou encore *il allait avoir fini* (qui nous semble pourtant grammaticale, bien qu'on puisse lui préférer *il aurait fini*).

Ensuite, nous ne voyons pas très bien ce qui justifie de distinguer **décalé** et **passé**. Pourquoi, par exemple, la vision ne peut-elle se combiner qu'aux verbes qui portent un grammème de temps? Pourquoi ne peut-elle pas apparaître seule? Ou encore, si le modèle de Martinet comprenait un grammème de temps **présent**, y aurait-il encore lieu de distinguer **décalé** et **passé**?

Autre caractéristique frappante de ce modèle: il va très clairement à l'encontre du postulat selon lequel la flexion est obligatoire (*cf.* Ch. IV, § 2, p. 51). Si on accepte ce postulat, on est forcé de voir des signes zéro¹⁰ là où il n'y a pas de marque formelle à l'intérieur d'un paradigme grammatical. Or, pour Martinet, le mot-forme (*je*) *sais* ne porte tout simplement ni mode ni temps; il ne s'agit pas de signes zéro, mais carrément d'une absence de signes grammaticaux. Cette absence contraste avec d'autres formes verbales qui, elles, portent des grammèmes de temps et de mode. Or, y a-t-il vraiment absence de signes ici? Sans trop entrer dans les détails (puisqu'on y reviendra plus loin, Ch. VI, § 3, p. 134 et Ch. VI, § 6, p. 177), on peut démontrer que le présent et l'indicatif correspondent bien à des signes. D'abord, en ce qui concerne le présent, s'il est vrai qu'il peut exprimer des sens très diversifiés, il reste qu'il ne peut pas signifier absolument n'importe quoi. Il existe un nombre limité de sens que peut exprimer cette forme et on peut les répertorier (nous le ferons plus loin, Ch. VI, § 6.4.4, p. 213). Pour l'indicatif, il est peut-être plus difficile de démontrer qu'il a un sens (nous en discuterons plus loin, Ch. VI, § 3, p. 134). Néanmoins, même en supposant qu'il n'ait pas de sens, on peut encore en identifier la combinatoire: l'indicatif peut se combiner à des morphèmes de temps, ce qui n'est pas le cas des autres modes (voir la discussion plus loin concernant le subjonctif imparfait). Ainsi, en tenant compte de la troisième composante des signes, à savoir leur combinatoire, on ne peut pas complètement exclure l'existence d'un grammème **indicatif**, quand bien même il aurait un signifiant nul et serait vide de sens, comme le pense Martinet. En fait, nous verrons au

10. À propos des signes zéro, voir notamment Mel'čuk (1997b: p. 21).

cours des prochaines pages, et surtout en discutant le modèle de Touratier à la section suivante, que le refus de considérer l'indicatif comme un mode mène à un modèle difficilement tenable.

C'est Menanteau (1986) qui nous mettra sur la piste. Il critiquait dans son article le modèle de Martinet, et nous reprenons ici quelques-unes de ses objections et ses conclusions.

D'abord, il soutient que la seule existence du subjonctif imparfait ne justifie pas de maintenir séparées les catégories de mode et de temps. Premièrement, le subjonctif imparfait est devenu extrêmement rare à l'oral, il est donc hasardeux de baser quoi que ce soit sur son existence. Ensuite, même si on admet que cette forme est d'usage suffisamment courant pour être considérée, Menanteau fait remarquer que, compte tenu du contexte où peut apparaître le subjonctif imparfait, les locuteurs ne perçoivent jamais de différence sémantique entre celui-ci et un subjonctif présent, la seule différence notable relevant du registre, ce qui laisse comprendre qu'au mieux on pourrait y voir un grammème indiquant le registre, mais pas un grammème de temps. Ainsi, aucun des modes de Martinet ne serait vraiment compatible avec un des temps (sauf peut-être dans *qu'il vienne de chanter*, mais continuons, nous verrons que cela n'est pas un problème).

Menanteau soutient également que les grammèmes **récent** et **prochain** ne peuvent pas faire partie de la catégorie de temps puisque cela mène à certaines contradictions. En effet, en considérant *vient de chanter* comme un temps de CHANTER, Martinet n'a d'autre choix pour rendre compte de *venait de chanter* que de stipuler une catégorie additionnelle, soit la vision, dont l'unique grammème (**décalé**) est associé aux mêmes sens et aux mêmes formes que le grammème de temps **passé**. Menanteau se questionne, à juste titre dirions-nous, sur le bien-fondé d'avoir deux grammèmes ayant le même sens et étant associés aux mêmes formes mais relevant de deux catégories séparées. La catégorie de vision serait donc un artefact qui proviendrait de l'impossibilité de combiner dans une même forme **récent** et **passé**, puisqu'ils font tous les deux partie d'une même catégorie flexionnelle. En sortant

récent de la catégorie de temps, le problème s'évapore. Le même raisonnement vaut pour **prochain**.

Menanteau propose de faire l'économie de la catégorie de vision. On se retrouve alors avec quatre modes et trois temps, tous mutuellement exclusifs, et qui ont tous la même combinatoire (abstraction faite des catégories d'accord). Suivant les principes de similitude de combinatoire et d'exclusion mutuelle établis par Martinet (*cf.* Ch. IV, § 6.1, p. 63), Menanteau conclut que les grammèmes de temps et de mode ne forment qu'une seule catégorie flexionnelle. Il obtient alors le modèle schématisé ci-dessous :

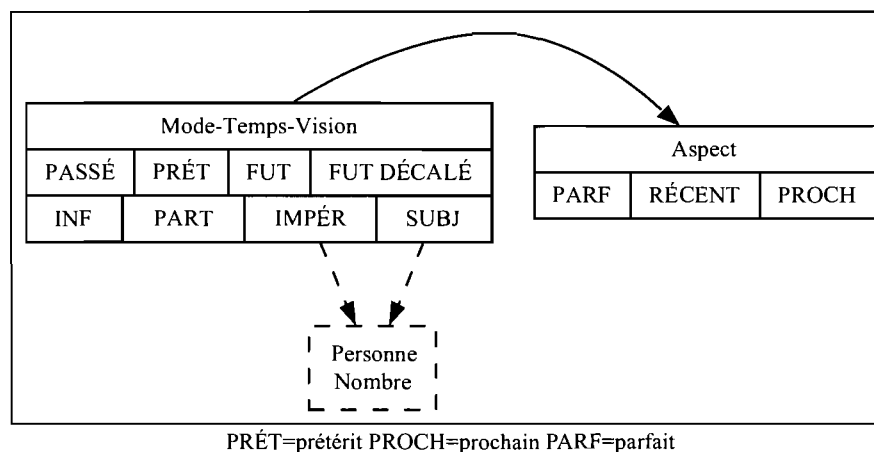


Figure 15 — Le modèle de Martinet revu par Menanteau

Touratier présentera et étoffera dix ans plus tard un modèle très similaire que nous discutons à la prochaine section, aussi gardons-nous certaines de nos objections pour plus tard. Pour l'instant, contentons-nous de noter que, même en acceptant que **récent** et **prochain** ne sont pas des temps comme le suggère Menanteau, la catégorie de vision de Martinet reste encore utile pour décrire le conditionnel, conçu par Martinet comme la combinaison du temps **futur** (–R–) et du grammème de vision **décalé** (–AI–). Si on souhaite éviter le problème de la duplication des grammèmes **passé/décalé**, il n'y a que deux solutions possibles. L'une est d'éliminer la catégorie de vision et de stipuler un grammème de temps supplémentaire, disons **futur décalé** : c'est ce que propose Menanteau. Mais on perd alors la compositionnalité du modèle de Martinet, qui tentait justement de rendre compte du fait

que le conditionnel à valeur temporelle est représentable, tant dans sa forme que dans son sens, en termes des signes –R– et –AI–. Or, nous voyons dans cette explication du conditionnel un des meilleurs éléments du modèle de Martinet, certainement pas une caractéristique à éliminer. L'autre solution envisageable consisterait à sortir le grammème **passé** de la catégorie de temps. C'est la voie empruntée par Touratier, que nous allons maintenant discuter.

3.2 Touratier

Dans *Le système verbal français* (1996), Touratier élabore un modèle construit sur les bases théoriques jetées par Martinet (1970, 1979). On peut en fait considérer ce modèle comme une version encore plus poussée de celui de Martinet.

Il faut prendre le système de Touratier comme un modèle des grammèmes superficiels, en nos termes (*cf.* Ch. IV, § 1, p. 46). En effet, il est fait mention dans le texte, par exemple, du futur périphrastique, et pourtant il ne figure nulle part en tant que grammème d'une catégorie flexionnelle. Puisqu'il est représentable dans son signifiant en termes de la combinaison «ALLER + V_{inf}», il n'a pas sa place dans un modèle des grammèmes superficiels.

Mises à part les catégories d'accord et la voix, il y a trois catégories flexionnelles chez Touratier: le mode, le temps et l'aspect (mais attention: la terminologie est trompeuse).

Les grammèmes de mode n'ont pas forcément un sens modal. En fait, la sémantique n'a à peu près aucune importance ici, puisque c'est essentiellement sur des critères basés sur la combinatoire que sont regroupés les grammèmes. On trouve dans la catégorie de mode des grammèmes aux sens les plus variés, allant du temporel au modal: **futur** (le suffixe –R–), **subjonctif**, **passé** (qui n'est pas celui de Martinet, mais qui correspond plutôt au passé simple), **impératif**, **infinitif** et **participe** (le même que chez Martinet). Tous sauf les deux derniers connaissent l'accord en personne et en nombre. En outre, les deux premiers (**futur** et **subjonctif**) se combinent à la catégorie de temps.

La catégorie de temps, selon Touratier, ne contient que le grammème **imparfait**, qui s'exprime parfois par le suffixe –AI– (qu'on retrouve tant à l'imparfait qu'au conditionnel) et parfois par le suffixe –ASS– (qu'on retrouve au subjonctif imparfait). Il correspond à la fois au **passé** et au **décalé** de Martinet. Son sens, bien que la catégorie à laquelle il appartient s'appelle «temps», ne doit pas être compris comme temporel. L'imparfait exprime un sens 'non actuel', dont l'interprétation varie selon le domaine auquel il est appliqué: temporel ou notionnel. Lorsqu'il est appliqué au domaine temporel, l'imparfait acquiert alors un sens temporel:

En réalité, quand son signifié est appliqué au temps, l'imparfait indique simplement que le procès concerné est 'non actuel au point de vue temporel', c'est-à-dire 'passé', et rien de plus. Et il se distingue alors du passé simple en ce qu'il ne précise pas que l'action inscrite dans le passé est objectivement délimitée et correspond à un fait historique bien isolable du reste du passé, avec un début et une fin parfaitement identifiés.

Touratier (1996: p. 110)

Les effets de sens de l'imparfait sont des plus variés quand il est appliqué au domaine notionnel (politesse, non réel, hypocoristique, etc).

Finalement, la catégorie d'aspect comprend deux grammèmes: **composé** (aussi appelé parfois «accompli») et **surcomposé**.

La Figure 16 et le Tableau X à la page suivante représentent de façon synthétique les catégories flexionnelles de Touratier, leurs grammèmes et la combinatoire de ceux-ci, avec un exemple pour chaque combinaison.

Les objections que nous avons quant à l'absence des grammèmes **indicatif**, **présent** et **inaccompli** chez Martinet valent également pour Touratier. Celui-ci se justifie, dans le cas du présent, en disant que ce grammème n'a pas de signifié propre, mais que les sens qu'on lui attribue proviennent du contexte:

Afin d'expliquer les différentes valeurs temporelles et non temporelles que les grammaires scolaires ont l'habitude de reconnaître au présent, nous dirons en effet qu'il s'agit des valeurs, non pas du verbe au présent lui-même, mais de l'énoncé dans lequel se trouve le verbe au présent.

Touratier (1996: p. 87)

V. Revue critique de quelques modèles de la conjugaison française

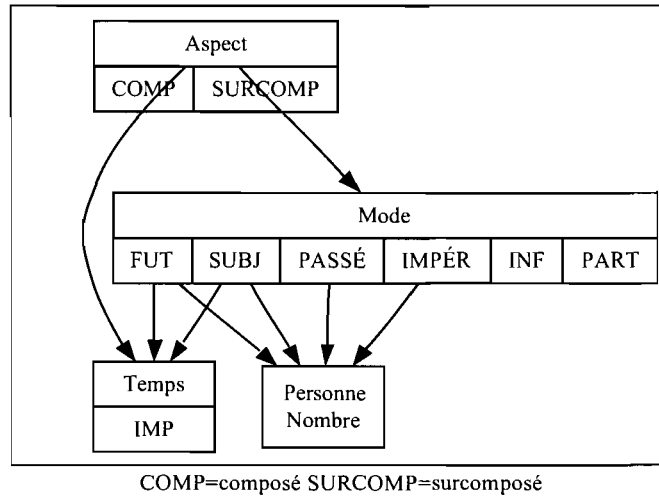


Figure 16 — Les catégories flexionnelles dans le modèle de Touratier

Mode	Temps	Aspect	Forme
infinitif			<i>savoir</i>
		composé	<i>avoir su</i>
		surcomposé	<i>avoir eu su</i>
participe			<i>sachant</i>
		composé	<i>ayant su</i>
		surcomposé	<i>ayant eu su</i>
impératif			<i>sache!</i>
		composé	<i>aie su!</i>
			sait
		composé	a su.
		surcomposé	a eu su
	imparfait		savait
		composé	avait su
		surcomposé	avait eu su

Tableau X — Le classement des formes selon Touratier

Mode	Temps	Aspect	Forme
subjonctif			sache
		composé	ait su
		surcomposé	ait eu su
	imparfait		<i>sût</i>
		composé	<i>eût su</i>
		surcomposé	<i>eût eu su</i>
passé			<i>sut</i>
		composé	<i>eut su</i>
		surcomposé	<i>eut eu su</i>
futur			saura
		composé	aura su
		surcomposé	aura eu su
	imparfait		saurait
		composé	aurait su
		surcomposé	aurait eu su

Tableau X — Le classement des formes selon Touratier

Cette position se heurte à au moins deux problèmes majeurs. D’abord, comment expliquer l’agrammaticalité de **Hier, je mange*? L’explication la plus simple n’est-elle pas qu’il y a contradiction entre le sens de *hier* et celui du présent? Ensuite, où est le contexte qui donne le sens d’habituel dans *Il travaille chez Bombardier*?

Il nous semble que cette conception s’appuie sur une vision du sens comme une sorte de point de gravité autour duquel tournent différentes «interprétations». Touratier rejette l’idée que le présent ait un sens par le fait que ce sens puisse «s’interpréter» de façons très diverses, voire contradictoires. Puisqu’il n’est pas possible de dégager de ces différentes interprétations une composante sémantique commune, et puisqu’il semble refuser implicitement l’hypothèse du caractère discret du sens, il conclut que le présent n’est porteur d’aucune signification. Le présent n’étant alors associé ni à une forme ni à un sens, il en arrive à la conclusion

que ce qu'on appelle «présent» n'est pas un signe. Nous avons dit plus haut, en parlant du modèle de Martinet, que si c'est vrai que le présent peut avoir différents sens, il ne peut pas signifier n'importe quoi (nous dressons plus loin, Ch. VI, § 6.4.4, p. 213, un inventaire partiel des sens de ce grammème). Par ailleurs, l'imparfait présente aussi une diversité de sens très similaire au présent (*cf.* Ch. VI, § 6.2, p. 193), et cela ne semble pas gêner Touratier. Bien sûr, comme l'imparfait a une forme bien identifiable, il serait plus difficile de prétendre qu'il n'existe pas.

Le grammème **imparfait** de Touratier, d'ailleurs, pose problème à nos yeux. La description que donne Touratier de son sens nous semble boiteuse; ses emplois temporels et notionnels n'ont rien à voir, si bien qu'on peut se demander s'il s'agit vraiment du même grammème. Mais ce qui nous gêne le plus par rapport à ce grammème **imparfait** est sa capacité, dans ce modèle, à se combiner à des grammèmes de «mode» *et* de commuter avec eux. Comparons les trois formes suivantes *il mangerait*, *qu'il mangeât* et *il mangeait*. Touratier y voit un grammème **imparfait** dans les trois cas. Mais dans les deux premiers, ce grammème se combine à un «mode», alors que dans la troisième forme, non. Encore une fois, il s'agit d'un effet du refus des fonctionnalistes du caractère obligatoire de la flexion (*cf.* Ch. IV, § 2, p. 51). Nous croyons qu'entre *mangeait* et *mangerait*, il y a une opposition entre le suffixe –R– et son absence. Cette absence en soi est à notre avis le signifiant d'un signe zéro qui manque dans le modèle de Touratier.

Touratier pousse le méthodologie de Martinet presque à l'extrême: tout repose sur la combinatoire des signes. C'est ainsi qu'il en vient à mettre dans une même catégorie des grammèmes qui n'ont rien à voir sémantiquement, mais qui présentent des traits de combinatoire similaires. Mais en fait, si on y regarde de plus près, on voit que l'auteur fait quelques entorses au principe selon lequel des grammèmes ayant une combinatoire semblable doivent être regroupés dans une même catégorie. Pour commencer, notons que seuls les grammèmes **futur** et **subjonctif** peuvent se combiner au grammème de temps **imparfait**¹¹. Ils n'ont donc pas la même

11. Rappelons que le grammème **imparfait** correspond à la fois au suffixe –AI– de l'indicatif imparfait ou du conditionnel et au suffixe –ASS– du subjonctif imparfait.

combinatoire que les autres grammèmes de leur catégorie. D'autre part, **infinitif** et **participe** ne sont pas compatibles avec l'accord en personne et en nombre. Ne faudrait-il donc pas postuler trois catégories au lieu d'une seule catégorie de mode : une première qui contienne les grammèmes **futur** et **subjonctif**, une seconde pour **passé** et **impératif**, et une troisième contenant **infinitif** et **participe**? Et encore, faut-il mettre l'impératif dans une classe à part puisqu'il est le seul à n'être compatible qu'avec certaines personnes, et non toutes? Quels sont les critères qui justifient de réunir tous ces grammèmes en une seule catégorie malgré leurs disparités? Jusqu'à quel point faut-il être flexible sur la similarité de combinatoire des grammèmes d'une même catégorie? Nous croyons que la combinatoire seule ne permet pas de trancher clairement: il faut tenir compte de la sémantique (*cf.* Ch. IV, § 6, p. 62).

4 Wilmet

Wilmet consacre un long chapitre de sa *Grammaire critique du français* (2003) à la flexion verbale. Sa terminologie bien particulière complique la comparaison de son modèle avec ceux issus d'autres cadres théoriques, mais nous préférons rester fidèle à sa nomenclature plutôt que d'en risquer une quelconque adaptation.

Le modèle de Wilmet contient, outre la voix et les catégories d'accord, trois catégories flexionnelles: mode, temps et aspect. Toutes les formes verbales expriment un grammème de chacune de ces catégories.

Wilmet se détache de la tradition notamment en ne reconnaissant pas l'impératif comme un mode :

L'«impératif» ne possède aucune forme en propre. Presque toutes se révèlent homophones de l'indicatif présent [...], quelques-unes du subjonctif. Deux métissent le subjonctif et l'indicatif¹².

Wilmet (2003 : p. 307)

12. Les deux formes métissées en question sont *sachons* et *sachez*, construites à partir d'un radical du subjonctif (*sach-*) et d'affixes de l'indicatif (*-ons / -ez*).

Le conditionnel, comme chez beaucoup de contemporains, est ramené à un temps de l'indicatif. Il lui reste donc comme formes à classer celles de l'infinif, de l'indicatif, du subjonctif et des deux participes. Il les divise dans un premier temps en formes impersonnelles (celles de l'infinif, du participe présent et du participe passé) et personnelles (celles de l'indicatif et du subjonctif) selon leur capacité à s'accorder avec un sujet, puis en formes actuelles (celles de l'indicatif) et inactuelles (celles du subjonctif et toutes les formes impersonnelles) selon leur capacité à situer les faits dans le temps. Le tout est ramené à une catégorie de mode qui contient trois grammèmes: **personnel actuel**, **personnel inactuel** et **impersonnel inactuel**. Les deux premiers renvoient respectivement aux formes de l'indicatif et du subjonctif. Le dernier mode se retrouve dans les formes de l'infinif et des deux participes, qui se distinguent par un grammème de la catégorie de temps.

Tous les grammèmes de temps se représentent en termes de relations temporelles entre un fait et un point de repère sur l'axe du temps. Il existe selon l'auteur trois types de repères par rapport auxquels les faits peuvent être mis en relation: un lieu, une personne ou une époque. La relation entre le fait et son repère ne peut être que l'une des trois suivantes: les faits peuvent lui être simultanés, antérieurs ou postérieurs. Il faut noter que les faits peuvent être mis en relation avec plus d'un repère à la fois (c'est le cas du conditionnel, qui met en relation un fait avec deux repères temporels). Enfin, les grammèmes principaux de temps peuvent être divisés en trois groupes, selon le repère auquel ils renvoient:

- Repère-lieu: **incident** (infinif), **décadent** (participe passé) et **incident-décadent** (participe présent), selon que le fait est «arrivant», «arrivé» ou «mi-arrivant, mi-arrivé» (ces termes ne sont pas définis explicitement dans le texte, mais ils semblent faire référence aux relations de précédence et de concomitance entre le fait et le repère-lieu mentionnées plus haut). Ces grammèmes servent à distinguer les formes du mode impersonnel inactuel.
- Repère-personne: **prospectif** (subjonctif présent) et **rétrospectif** (subjonctif imparfait), selon que le fait «épouse la ligne du temps» ou qu'il «la

prenne à rebours» (encore une fois, nous n'avons pas trouvé la définition de ces expressions, mais d'après les figures il semble que **prospectif** situe le fait après le repère-personne, et **rétrospectif**, avant celui-ci). Ces grammèmes distinguent les formes du mode personnel inactuel (c'est-à-dire le subjonctif).

- Repère-époque: **présent** (pour les formes du présent), **passé** (pour les formes de l'imparfait et du passé simple), **futur** (pour le futur simple) et **futur du passé** (pour le conditionnel, qui implique deux repères temporels). Le sens de ces grammèmes est bien celui que laissent supposer leurs noms. Ces grammèmes servent à distinguer les formes du mode personnel actuel (c'est-à-dire l'indicatif).

Le tout est multiplié par trois pour rendre compte, en plus des formes simples, des formes composées et surcomposées, ce qui donne les grammèmes **incident antérieur** (infinitif passé), **incident bisantérieur** (infinitif surcomposé), **prospectif antérieur** (subjonctif passé — *ait su*), **prospectif bisantérieur** (subjonctif surcomposé — *ait eu su*), etc. Seul le «décadent bisantérieur» (qui correspondrait à un participe passé surcomposé — **eu eu su*) n'existe pas, ce qui laisse en tout 26 grammèmes de temps (*cf.* Figure 17 ci-dessous).

Enfin, les grammèmes d'aspect indiquent si le repère se trouve à l'intérieur des bornes du fait (c'est le sens du grammème **sécant**) ou à l'extérieur de celles-ci (c'est le sens du grammème **global**). Tout comme les temps, les aspects sont multipliés par trois pour rendre compte encore une fois des formes composées et surcomposées, ce qui donne, en plus des deux grammèmes déjà mentionnés, les grammèmes **sécant extensif**, **sécant bisextensif**, **global extensif** et **global bisextensif**¹³.

La figure suivante représente sous forme graphique les catégories flexionnelles du modèle de Wilmet ainsi que leur combinatoire. Il est à noter que toutes les com-

13. Les grammèmes d'aspect ne correspondent pas à des formes facilement identifiables. Le lecteur pourra se référer au Tableau XI ci-dessous pour des exemples de formes qui expriment ces grammèmes.

binaisons ne sont pas permises et que les incompatibilités mineures ne sont pas indiquées dans cette figure. Par contre, le Tableau XI ci-dessous donne la liste de toutes les combinaisons possibles, avec un exemple pour chacune.

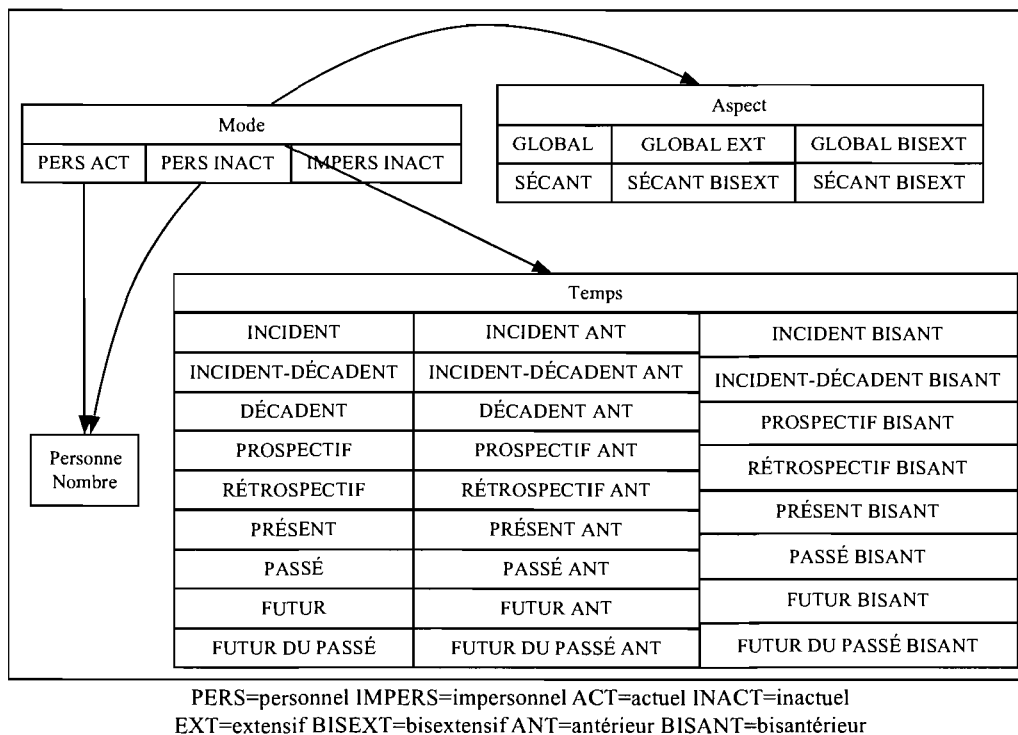


Figure 17 — Les catégories flexionnelles dans le modèle de Wilmet

Mode	Temps	Aspect	Forme
impersonnel inactuel	incident	global	<i>savoir</i>
	incident-décadent	sécant	<i>sachant</i>
	décadent	global	<i>su</i>
	incident antérieur	global extensif	avoir su
	incident-décadent antérieur	sécant extensif	ayant su
	décadent antérieur	global extensif	<i>eu su</i>
	incident bisantérieur	global bisextensif	ayant eu marché
	incident-décadent bisantérieur	sécant bisextensif	ayant eu marché
personnel inactuel	prospectif	global	<i>sache</i>
	rétrospectif	global	<i>sût</i>
	prospectif antérieur	global extensif	<i>ait su</i>
	rétrospectif antérieur	global extensif	<i>eût su</i>
	prospectif bisantérieur	global bisextensif	<i>ait eu su</i>
	rétrospectif bisantérieur	global bisextensif	<i>eût eu su</i>

Tableau XI — Le classement des formes selon Wilmet

Mode	Temps	Aspect	Forme
personnel actuel	présent	sécant	<i>sait</i>
	passé	global	<i>sut</i>
		sécant	<i>savait</i>
	futur	global	<i>saura</i>
	futur du passé	global	<i>saurait</i>
	présent antérieur	sécant extensif	<i>a su</i>
	passé antérieur	global extensif	<i>eut su</i>
		sécant extensif	<i>avait su</i>
	futur antérieur	global extensif	<i>aura su</i>
	futur du passé antérieur	global extensif	<i>aurait su</i>
	présent bisantérieur	sécant bisextensif	<i>a eu su</i>
	passé bisantérieur	global bisextensif	<i>eut eu su</i>
		sécant bisextensif	<i>avait eu su</i>
	futur bisantérieur	global bisextensif	<i>aura eu su</i>
futur du passé bisantérieur	global bisextensif	<i>aurait eu su</i>	

Tableau XI — Le classement des formes selon Wilmet

Ce modèle nous semble peu économique étant donné le nombre élevé de grammèmes qu'il suppose. Les trois catégories flexionnelles de mode, temps et aspect suffisent en principe à décrire jusqu'à 624 formes verbales différentes¹⁴, alors que l'auteur ne cherche qu'à en décrire 29. En fait, la catégorie flexionnelle de temps à elle seule, avec ses 26 grammèmes, est presque suffisante pour décrire l'ensemble des phénomènes considérés. En lui ajoutant seulement trois grammèmes, on pourrait faire l'économie des deux autres catégories flexionnelles pour obtenir un modèle plat ne contenant que des grammèmes cumulatifs.

14. 4 modes × 6 aspects × 26 temps = 624 formes.

Un corollaire de cela est que chez Wilmet, les grammèmes n'ont pas toujours de valeur distinctive. Notamment, on peut douter de l'utilité de la catégorie d'aspect dans la plupart des cas, puisque dès que l'on connaît le grammème de temps on peut prédire quel sera le grammème d'aspect. Le seul cas où cela n'est pas possible est lorsque le grammème de temps est **passé** (l'aspect permet alors de différencier l'imparfait du passé simple). On peut donc se demander si les grammèmes d'aspect stipulés pour les autres formes sont bien des grammèmes ou s'il ne s'agit pas plutôt d'une composante du sens des grammèmes de temps.

Le modèle de Wilmet a néanmoins le mérite de faire une distinction relativement nette entre les «modes personnels» (indicatif et subjonctif) et les «modes impersonnels» (infinitif, participe présent et participe passé). Il y a un très large consensus sur cette division (sans doute parce qu'elle est basée sur des faits directement observables) et pourtant, peu de grammairiens organisent leur modèle en conséquence. Nous aimerions cependant voir ces deux groupes encore plus clairement divisés, c'est-à-dire dans des catégories flexionnelles différentes. Cela permettrait notamment de représenter de façon plus économique la combinatoire des grammèmes. On pourrait alors dire, par exemple, que tout ce qui est « personnel » a besoin des grammèmes d'accord, ce qui n'est pas possible dans le modèle de Wilmet tel quel, où il faut dire que **personnel actuel** et **personnel inactuel** se combinent aux grammèmes d'accord; ces deux grammèmes ne forment pas une classe à laquelle on pourrait faire référence dans son ensemble.

Cependant, nous nous questionnons sur la pertinence de la différence temporelle stipulée par Wilmet entre les formes de l'infinitif et des participes. Si à la rigueur nous pouvons concevoir qu'il y ait une différence d'ordre temporelle entre les deux participes, il nous semble douteux d'affirmer que l'infinitif porte un quelconque sens temporel. Quand l'auteur dit que l'infinitif exprime (nous soulignons) «le potentiel inentamé du temps *incident* et de l'aspect *global*, qui laissent aux procès leur potentialité maximale» (p. 350), ne veut-il pas dire justement que l'infinitif ne situe tout simplement pas les faits dans le temps ?

D'ailleurs, la caractérisation sémantique des temps de Wilmet nous laisse perplexe. Nous comprenons mal comment un fait peut être dans une relation *temporelle* avec un lieu ou une personne. Et tout d'abord, de quel lieu et de quelle personne s'agit-il? Nous supposons que l'auteur fait référence au «ici» et au «moi» de la situation d'énonciation, mais nous n'en avons pas trouvé la confirmation dans le texte.

D'autre part, la Figure 17 montre bien que les grammèmes de temps et d'aspect peuvent être divisés en trois groupes: «Ø», «antérieur» et «bisantérieur» pour les temps et «Ø», «extensif» et «bisextensif» pour les aspects. Qui plus est, on peut voir au Tableau XI que ces séries sont toujours parfaitement synchronisées et qu'elles sont directement liées à la présence de l'auxiliaire AVOIR: les formes sans auxiliaire sont toutes ni «antérieures» ni «extensives», celles avec un auxiliaire sont toutes à la fois «antérieures» et «extensives», et celles avec deux auxiliaires sont toutes aussi «bisantérieures» que «bisextensives». Il est parfaitement clair que «antérieur» et «extensif» dans ce modèle décrivent des composantes sémantiques de l'auxiliaire AVOIR. Alors pourquoi ne pas en avoir fait une catégorie flexionnelle? Avec une catégorie flexionnelle supplémentaire, disons de «composition» pour rester neutre, contenant trois grammèmes (**simple, composé, surcomposé**), on aurait réduit le nombre d'aspects à deux (**sécant et global**), et le nombre de temps à neuf (**incident, incident-décadent, décadent, prospectif, rétrospectif, présent, passé, futur et futur du passé**).

Si on fait abstraction des temps «antérieurs» et «bisantérieurs», parmi les neuf temps restants on peut distinguer encore trois séries indépendantes. En effet, tous ne se combinent pas aux mêmes grammèmes de mode: le mode **personnel actuel** ne se combine qu'aux temps à repère temporel (**présent, passé, futur et futur du passé**), le mode **personnel inactuel** se combine aux temps à repère-personne (**prospectif et rétrospectif**), alors que le mode **impersonnel inactuel** se combine aux temps à repère-lieu (**incident, incident-décadent et décadent**). Il y a donc en fait chez Wilmet trois catégories flexionnelles de temps, selon le type de point de repère par rapport auquel sont situés les faits.

On peut essayer de revoir le modèle de Wilmet en divisant ce qui peut l'être. On obtient alors des catégories flexionnelles contenant moins de grammèmes, comme le montrent la Figure 18 et le Tableau XII ci-dessous.

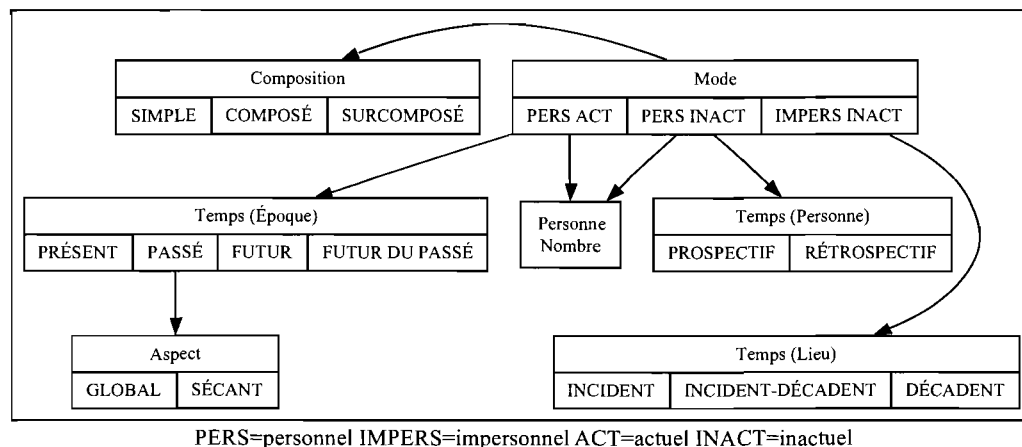


Figure 18 — Les catégories flexionnelles dans le modèle de Wilmet révisées

Mode	Temps (Lieu)	Temps (Personne)	Temps (Époque)	Aspect	Forme simple ~ composée ~ surcomposée
impersonnel inactuel	incident				<i>savoir</i> avoir su —
	incident-décadent				<i>sachant</i> ayant su —
	décadent				su eu su —
personnel inactuel		prospectif			<i>sache</i> ait su ait eu su
		rétrospectif			<i>sût</i> eût su eût eu su

Tableau XII — Le classement des formes selon Wilmet révisé

Mode	Temps (Lieu)	Temps (Personne)	Temps (Époque)	Aspect	Forme simple ~ composée ~ surcomposée
personnel actuel			présent		<i>sait</i> <i>a su</i> <i>a eu su</i>
			passé	global	<i>sut</i> <i>eut su</i> <i>eut eu su</i>
				sécant	<i>savait</i> <i>avait su</i> <i>avait eu su</i>
			futur		<i>saura</i> <i>aura su</i> <i>aura eu su</i>
			futur du passé		<i>saurait</i> <i>aurait su</i> <i>aurait eu su</i>

Tableau XII — Le classement des formes selon Wilmet revisité

Remarquons la présence surprenante de **eu su* dans les formes prévues par le modèle de Wilmet. De l'aveu de l'auteur, cette forme ne peut pas être utilisée «à l'état libre». De ses justifications, la plus intéressante est que les formes surcomposées des constructions pronominales du type *Pierre s'est eu lavé* n'admettraient que le découpage *s'est + eu lavé*. En fait, cela indique surtout que l'aspect exprimé par AVOIR forme une nouvelle base sur laquelle s'applique la voix, un point de vue que nous défendons plus loin (Ch. VI, § 5, p. 164).

Enfin, nous ne sommes pas d'accord avec les jugements de grammaticalité de Wilmet sur les formes surcomposées. Il nous semble que les seules formes surcomposées réellement utilisées en français standard contemporain sont celles construites sur le présent ou l'imparfait (*a eu su* et *avait eu su*). Les autres formes surcomposées, comme *avoir eu su* ou *ayant eu su*, ne se trouvent guère ailleurs que dans les grammaires et dans quelques dialectes régionaux (voir la discussion à ce sujet Ch. II, § 4, p. 31). Par ailleurs, notons que Wilmet prend en compte les for-

mes de l'imparfait du subjonctif, que nous avons écartées de notre étude (cf. Ch. II, § 2, p. 27).

5 Gosselin

Dans *Sémantique de la temporalité en français* (1996), Gosselin construit un modèle formel du temps et de l'aspect en français. Il cherche à construire un modèle computationnel capable d'identifier, en analyse, uniquement les relations temporelles et aspectuelles en jeu dans un texte. Pour cette raison, il s'intéresse essentiellement aux formes de l'indicatif¹⁵ et ne dit rien à propos du mode, de la voix, etc., contrairement aux autres auteurs présentés ici. Par ailleurs, son modèle n'en est pas un composé de grammèmes et de catégories flexionnelles au sens ou nous l'entendons, mais il se laisse assez facilement représenter sous cette forme.

Son modèle s'appuie sur un formalisme simple où des intervalles sont disposés sur un axe du temps. Ce sont les relations entre ces intervalles (ou les bornes qui les délimitent) qui décrivent les sens temporels et aspectuels de la langue. Il y a quatre types d'intervalles :

- L'intervalle $[01,02]$ ¹⁶ correspond à la durée de l'énonciation. C'est le principal point d'ancrage temporel et il n'y en a qu'un seul par phrase.
- L'intervalle $[B1,B2]$ correspond à la durée du fait dénoté par le verbe. Il y en a au moins un par proposition¹⁷.
- L'intervalle $[I,II]$ est l'intervalle de référence par rapport auquel est situé le fait dénoté par le verbe. Il y en a au moins un par proposition¹⁸.
- Les intervalles $[ct1,ct2]$ correspondent aux circonstanciels (s'il y en a).

15. Il glisse également quelques mots au sujet des subjonctifs présent et imparfait, mais nous laisserons de côté ces formes ici.

16. On réfère aux intervalles par les bornes qui les délimitent: l'intervalle $[i,j]$ correspond à la section de l'axe du temps comprise entre l'instant i et l'instant j .

17. Les formes composées ont en effet deux intervalles $[B1,B2]$, nous le verrons plus loin.

18. Idem.

Les relations possibles entre ces intervalles sont au nombre de huit :

- L'antériorité: « $[i,j]$ ANT $[k,l]$ » signifie que l'intervalle $[i,j]$ est entièrement avant la borne k .
- La postériorité: « $[i,j]$ POST $[k,l]$ » signifie que l'intervalle $[i,j]$ est entièrement après la borne l .
- La succession: « $[i,j]$ SUCC $[k,l]$ » signifie que $[i,j]$ commence après k (mais $[i,j]$ peut être à l'intérieur de $[k,l]$).
- La précédençe: « $[i,j]$ PREC $[k,l]$ » signifie que $[i,j]$ commence avant k (mais $[k,l]$ peut être à l'intérieur de $[i,j]$).
- La simultanéité: « $[i,j]$ SIMUL $[k,l]$ » signifie que l'intersection entre les deux intervalles n'est pas vide (les intervalles peuvent coïncider ou se chevaucher).
- La coïncidence: « $[i,j]$ CO $[k,l]$ » signifie que les bornes de $[i,j]$ et de $[k,l]$ coïncident exactement.
- Le recouvrement: « $[i,j]$ RE $[k,l]$ » signifie que $[k,l]$ est entièrement contenu dans $[i,j]$ et que les bornes ne coïncident pas.
- L'accessibilité: « $[i,j]$ ACCESS $[k,l]$ » signifie que $[k,l]$ est entièrement contenu dans $[i,j]$, et les bornes peuvent coïncider.

Ce sont ces relations qu'expriment les grammèmes de temps et d'aspect. Les catégories flexionnelles de Gosselin se définissent selon la nature des intervalles qui sont en jeu. Il y a trois catégories: aspect, temps absolu et temps relatif.

Les grammèmes de la catégorie d'aspect expriment des relations entre l'intervalle de référence ($[I,II]$) et l'intervalle du fait dénoté par le verbe ($[B1,B2]$). Il y a selon Gosselin quatre relations de ce type en français, qui correspondent à quatre aspects différents: **aoristique** ($[I,II]$ CO $[B1,B2]$, qui ne s'exprime que par le passé simple et donne une «vue globale» du fait), **inaccompli** ($[B1,B2]$ RE $[I,II]$, qu'on trouve à l'imparfait ou au présent et qui donne une «vue partielle» du fait), **accompli** ($[I,II]$ POST $[B1,B2]$, qui s'exprime par l'auxiliaire AVOIR <ÊTRE> et qui fait valoir «l'état résultant» du fait) et **prospectif** ($[I,II]$ ANT $[B1,B2]$, qui s'exprime par l'auxiliaire ALLER et qui montre «l'état préparatoire»).

Les grammèmes de la catégorie de temps absolu expriment des relations entre l'intervalle de référence ([I,II]) et le moment d'énonciation ([01,02]). Gosselin rompt avec la définition généralement acceptée du temps absolu (*cf.* notamment Comrie 1985), selon laquelle c'est le fait dénoté par le verbe qui est directement mis en relation avec le moment d'énonciation (c'est-à-dire que la relation est entre [B1,B2] et [01,02]). Cet écart vise à rendre compte du fait que dans une phrase comme *Quand j'ai regardé par la fenêtre, le petit dormait*, rien n'indique si le petit dort encore au moment où je parle ou non, donc l'imparfait ne situe pas le fait directement par rapport au moment d'énonciation selon Gosselin. Il y a trois grammèmes de temps absolu dans son modèle: **présent** ([I,II] SIMUL [01,02])¹⁹, **passé** ([I,II] ANT [01,02]), qu'on trouve dans l'imparfait comme dans le passé simple) et **futur** ([I,II] POST [01,02]).

Les grammèmes de la catégorie de temps relatif expriment des relations entre l'intervalle de référence d'une première proposition ([I,II]) et celui d'une seconde proposition qui lui est subordonnée ([I',II']). Encore une fois, Gosselin s'écarte de la définition la plus courante du temps relatif (*cf.* encore une fois Comrie 1985), selon laquelle c'est le fait dénoté et l'intervalle de référence qui sont mis en relation (nous verrons un peu plus loin, p. 112, les raisons de cet écart). Il y a trois grammèmes de temps relatif dans son modèle: **simultané** ([I',II'] SIMUL [I,II]), **antérieur** ([I',II'] ANT [I,II]), qu'on trouve à l'imparfait comme au passé simple) et **ultérieur** ([I',II'] POST [I,II]), qui s'exprime par le conditionnel).

Les deux catégories de temps ne se combinent pas entre elles. Elles peuvent par contre se combiner à la catégorie d'aspect. Ainsi, le présent exprime le grammème de temps absolu **présent** et l'aspect **inaccompli**, alors que le passé simple exprime à la fois le temps absolu **passé** et l'aspect **oristique**.

Les formes composées sont conçues comme la combinaison de deux verbes qui ouvrent chacun ses propres intervalles:

19. Le grammème **présent** est défini par l'auteur via la relation SIMUL, mais il note par ailleurs que la simultanéité stricte entre le fait et le moment d'énonciation est rare, et qu'il s'agit le plus souvent de la relation CO.

Un temps composé est virtuellement porteur de deux relations temporelles et de deux aspects distincts. Plus précisément, on fait l'hypothèse que le participe passé présente toujours (sauf avec le passif) le procès sous un aspect aoristique et une relation temporelle relative d'antériorité par rapport à l'intervalle de référence associé à l'auxiliaire; alors que celui-ci marque un temps absolu et un aspect inaccompli (au moins au plus-que-parfait et au passé composé) sur l'état résultant du procès.

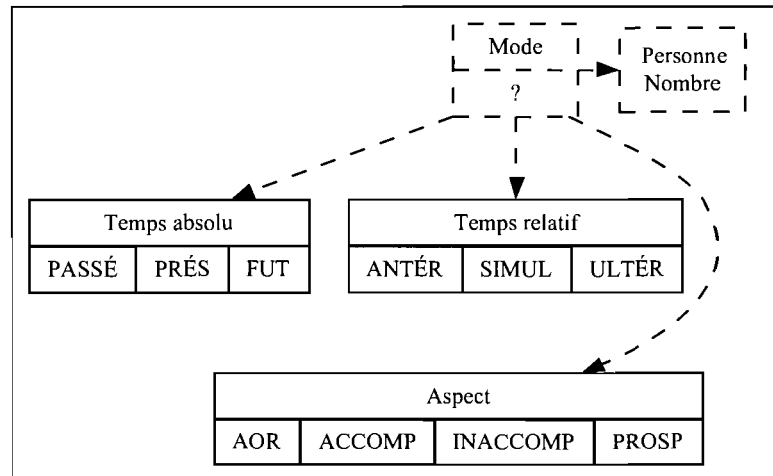
Gosselin (1996: p. 145).

Le participe passé correspond au fait lui-même ([B1,B2]) alors que l'auxiliaire exprime la situation résultante de ce procès ([B1',B2']). Ainsi dans le passé composé, le participe passé exprime les grammèmes **antérieur** (l'intervalle de référence du procès est avant l'intervalle de référence de l'auxiliaire, c'est-à-dire celui de la proposition) et **aoristique** (le fait dénoté par le participe passé est pris dans son ensemble), alors que l'auxiliaire exprime les grammèmes **présent** (la situation résultante est actuelle) et **inaccompli** (l'état résultant est vu partiellement)²⁰. La différence entre le passé composé qui exprime un passé et celui qui exprime un accompli tient à la saillance communicative de [B1,B2] (le fait dénoté par le participe passé) par rapport à [B1',B2'] (la situation résultante dénotée par l'auxiliaire): si [B1,B2] est plus saillant, alors c'est un passé, sinon c'est un présent accompli. C'est le contexte qui permet de déterminer lequel est le plus saillant.

La Figure 19 et le Tableau XIII ci-dessous résument les catégories flexionnelles du modèle de Gosselin ainsi que leurs grammèmes et la combinatoire de ceux-ci, avec des exemples pour les différentes combinaisons possibles. Pour les formes composées nous indiquons dans le Tableau XIII, p. 110, les grammèmes portés par l'auxiliaire (Aux) et ceux portés par le participe passé (PP).

À notre avis, le traitement des formes composées dans ce modèle est problématique. Tout d'abord, il ne permet pas de décrire les formes surcomposées (à propos desquelles nous n'avons pas trouvé de commentaire dans cet ouvrage). De plus, comment se fait-il que le participe passé n'exprime les grammèmes **antérieur** et **aoristique** que quand il est utilisé avec un auxiliaire, mais pas dans les formes

20. Plus loin, lorsque Gosselin donne les informations sémantiques portées par chaque forme verbale, il ne mentionne pas l'aspect inaccompli pour le passé composé.



ANTÉR=antérieur SIMUL=simultané ULTÉR=ultérieur AOR=oristique
 ACCOMP=accompli INACCOMP=inaccompli PROSP=prospectif

Figure 19 — Les catégories flexionnelles dans le modèle de Gosselin

Temps absolu	Temps relatif	Aspect	Forme
présent		inaccompli	sait
passé		oristique	<i>su</i>
		inaccompli	savait
futur		inaccompli	saura
	ultérieur	inaccompli	<i>sauraitⁱ</i>
Aux: présent	PP: antérieur	PP: oristique Aux: inaccompli	<i>a su</i>
Aux: passé	PP: antérieur	PP: oristique Aux: oristique	<i>eut su</i>
		PP: oristique Aux: inaccompli	<i>avait su</i>
Aux: futur	PP: antérieur	PP: oristique Aux: inaccompli	<i>aura su</i>
	PP: antérieur Aux: ultérieur	PP: oristique Aux: inaccompli	<i>aurait suⁱ</i>
?		?	<i>va savoirⁱⁱ</i>
?		?	<i>allait savoirⁱⁱ</i>

Tableau XIII — Le classement des formes selon Gosselin

- i. Le conditionnel et le conditionnel passé ne peuvent être mis en relation qu'avec un intervalle de référence appartenant à un verbe au passé.

V. Revue critique de quelques modèles de la conjugaison française

- ii. Les formes avec ALLER n'apparaissent pas dans le chapitre où Gosselin donne le sens de chaque forme verbale, si bien que nous ne savons pas comment les décrire.

passives? Comment peut-on être sûr que ces sens ne font pas plutôt partie du sens de AVOIR?

En outre, il y a une certaine confusion entre grammèmes profonds et grammèmes de surface qui entraîne des contradictions, du moins en apparence. Il semble qu'il faille comprendre les quatre grammèmes d'aspect mentionnés plus haut comme des grammèmes profonds, puisque les formes avec AVOIR ne sont pas décrites comme portant le grammème **accompli** (probablement parce que ce dernier est exprimé, selon Gosselin, par l'auxiliaire lui-même). Mais alors, comment traiter *avait mangé*? Si l'auxiliaire lui-même exprime le grammème **accompli**, peut-il en plus porter le grammème **inaccompli** (cf. l'imparfait sur l'auxiliaire)? On a le même problème avec l'auxiliaire ALLER.

Par ailleurs, nous regrettons que l'auteur ne dise rien sur les formes composées aux autres modes qu'à l'indicatif comme *avoir mangé*, *ayant mangé*, etc. Ces formes ne portent-elles pas de sens aspectuel? Peut-être ne faut-il pas voir l'accompli et le prospectif comme des grammèmes, mais simplement comme des étiquettes pour désigner les auxiliaires AVOIR et ALLER?

Toutefois, nous trouvons particulièrement intéressante l'idée de redéfinir les grammèmes de temps absolu en termes de relations entre le moment d'énonciation et le moment de référence (et non le moment du fait dénoté par le verbe). Cela ouvre la porte à un système temporel articulé, où une catégorie situe un moment de référence par rapport à 'maintenant' alors qu'une autre situe les faits par rapport à ce moment de référence. C'est précisément le genre de système que nous tenterons de développer (Ch. VI, § 6, p. 177). Cependant, pour obtenir un tel système, il faudrait une catégorie qui situe les faits par rapport au moment de référence. Chez Gosselin, les relations [B1,B2] / [I,II] sont des relations aspectuelles et non temporelles. Les temps relatifs, quand à eux, ne disent rien quant au rapport [B1,B2] / [I,II] et sont de toute façon en concurrence avec les temps absolus.

À propos des temps relatifs, nous ne sommes pas convaincu qu'ils mettent en relation deux intervalles de référence (plutôt que deux faits). Gosselin justifie sa définition du temps relatif en se basant sur la phrase *Pierre disait que, lundi, Luc aurait terminé son travail depuis longtemps* (Gosselin, 1996: p. 21). Il soutient que cette phrase ne situe pas l'intervalle de référence de la subordonnée par rapport au moment d'énonciation, c'est-à-dire qu'elle n'établit pas de relation $[I',II'] / [01,02]$. En effet, rien ne dit si le lundi en question est dans le passé ou dans le futur. Il soutient également que cette phrase n'implique aucune relation temporelle entre 'dire' et 'terminer', c'est-à-dire entre ce qu'il considère comme les faits de la principale et de la subordonnée, donc qu'elle n'établit pas de relation $[B1',B2'] / [B1,B2]$. Il est vrai que dans cet exemple, Luc peut très bien avoir déjà terminé son travail au moment où Pierre parle (et il sera d'autant plus vrai dans le futur que le travail soit terminé depuis longtemps). Cependant, nous croyons que cet exemple porte à confusion puisque le fait dénoté dans la subordonnée est selon nous 'avoir terminé', et non 'terminer', et c'est sur ce sens que porte celui du conditionnel²¹. Le problème est que 'avoir terminé' est un état, et non un changement d'état (contrairement à 'terminer'). Or, dire qu'un état existera à un moment donné n'implique en rien qu'il ne puisse exister à d'autres moments: quand on dit *Demain, Paul aura les cheveux noirs*, cela ne signifie pas que les cheveux de Paul ne sont pas déjà noirs aujourd'hui. Donc, le fait que l'on ne sache pas où 'terminer' se situe par rapport à 'dire' dans l'exemple de Gosselin n'est pas pertinent, selon nous. Ce qui est important est qu'il est dit explicitement que 'avoir terminé' vient après 'dire'; tout ce qui n'est pas dit, c'est le moment où le changement d'état se produit. Aussi croyons-nous que les temps relatifs mettent directement en relation les faits de la subordonnée et de la principale ($[B1',B2'] / [B1,B2]$) plutôt que leurs intervalles de références ($[I',II'] / [I,II]$).

Le formalisme de Gosselin a l'avantage non négligeable d'être simple et efficace. Il encode parfaitement bien les relations temporelles et semble relativement facile à implémenter dans un système de traitement de la langue. Cependant, il a

21. Voir à ce propos notre discussion de l'aspect accompli, Ch. VI, § 5, p. 164.

ses limites. Nous avons déjà critiqué ce formalisme au chapitre précédent (Ch. IV, § 5, p. 60); nous résumons brièvement nos objections ici. D'abord, ce formalisme ne permet pas de représenter les acceptions non-temporelles des grammèmes. Il n'est donc utile que pour décrire une partie du système flexionnel, ce qui s'insère mal dans un modèle global de la langue. Ensuite, ce type de représentation ne permet pas de tenir compte de la perspective que peuvent encoder les signes grammaticaux. Par exemple, l'accompli se définit par la relation [I,II] POST [B1,B2] (l'intervalle de référence vient après le fait dénoté par le verbe). Cette relation est strictement équivalente à [B1,B2] ANT [I,II]. On ne peut donc pas encoder la différence entre 'le fait est avant l'intervalle de référence' et 'l'intervalle de référence est après le fait'. Or, nous croyons que cette différence communicative est pertinente en français (nous y reviendrons au chapitre suivant, § 5.2, p. 166 et § 6.4.2, p. 205).

Finalement, le modèle de Gosselin présente deux catégories flexionnelles (temps absolu et temps relatif) en compétition. Nous avons déjà exprimé nos réticences par rapport à ce type de modèle lors de notre critique du modèle traditionnel (§ 1, p. 72).

6 Synthèse

Nous avons présenté et critiqué quelques modèles proposés par d'autres auteurs. Nous reprenons ci-dessous nos principales critiques.

L'opposition traditionnelle entre modes personnels et modes impersonnels cache à notre avis quelque chose de plus profond qu'une simple différence de combinatoire. Les différences sémantiques importantes entre ces deux groupes nous portent à croire qu'il s'agit en fait de deux catégories flexionnelles (ce que nous développons au chapitre suivant, Ch. VI, § 1, p. 116).

Les modèles plats où le temps et l'aspect ne sont pas distingués ne rendent pas compte de la régularité tant sémantique que formelle de la composition avec

l'auxiliaire AVOIR <ÊTRE>. Cependant, faire appel à une catégorie d'aspect pour distinguer les trois principales formes du passé (passé simple, imparfait et passé composé) pose problème dans le traitement des formes composées et surcomposées. D'ailleurs, dans les modèles où l'aspect est une catégorie distincte du temps, on remarque une certaine confusion quant aux deux valeurs du passé composé (qui peut marquer autant l'antériorité que l'accomplissement). Nous croyons que de voir dans le passé composé la forme de deux grammèmes distincts permettrait de mieux représenter le temps et l'aspect en français, et par la même occasion résoudre l'énigme des formes surcomposées, souvent laissées de côté (nous développerons ce point au prochain chapitre, Ch. VI, § 4, p. 144). Nous croyons également que ce qui distingue les trois principaux temps du passé n'est pas l'aspect, mais qu'il s'agit d'une illusion créée par le jeu de deux catégories de nature temporelle (ce que nous défendons plus loin Ch. VI, § 6, p. 177).

Bon nombre d'auteurs s'intéressent surtout à la description du sens des différentes formes verbales. Nous avons vu que les méthodologies basées uniquement soit sur le sens, soit sur la forme, soit sur la combinatoire des grammèmes mènent à des modèles qui ne représentent pas un système de signes, mais simplement un système sémantique, formel ou combinatoire. Nous croyons qu'il faut toujours prendre en compte les trois composantes des grammèmes lors de l'élaboration d'un modèle (nous avons déjà présenté notre méthodologie au chapitre précédent, Ch. IV, § 6, p. 62).

La distinction entre temps absolus et temps relatifs est généralement présentée soit comme deux sous-ensembles de la catégorie de temps, soit comme deux catégories de temps concurrentes. Il nous semble qu'une troisième voie s'impose, soit de postuler une catégorie de décalage qui se combine aux temps (nous défendons ce point de vue au chapitre suivant, Ch. VI, § 6, p. 177).

Nous allons maintenant proposer notre propre modèle en tâchant de résoudre ces problèmes.

VI. Un modèle descriptif de la conjugaison française

Nous avons vu au chapitre précédent que les modèles de la conjugaison française de nos prédécesseurs ne se laissaient pas représenter de façon satisfaisante sous forme d'un système flexionnel répondant aux critères définis au Chapitre IV (c'est-à-dire un système où les grammèmes sont regroupés en catégories flexionnelles quand ils sont mutuellement exclusifs, ils ont une combinatoire similaire et ils présentent une parenté sémantique évidente dans leur acception de base). Dans le présent chapitre, nous présentons une description systématique du temps grammatical en français, en suivant la méthodologie décrite au Chapitre IV, afin de construire un système cohérent qui puisse mettre en correspondance l'ensemble des formes exprimant le temps grammatical (y compris les formes surcomposées, souvent laissées de côté dans les grammaires) avec les sens qu'elles expriment (et vice-versa). De façon secondaire, nous nous intéresserons également à l'aspect tel qu'exprimé dans les formes verbales françaises.

Nous ne nous attarderons pas sur les problèmes de la voix, du mode et de l'accord verbal. Nous nous pencherons surtout sur les formes de l'indicatif¹, puisque ce sont elles qui expriment les grammèmes de temps et d'aspect qui nous intéressent. Cependant, les notions de temps, d'aspect, de mode et de finitude sont si étroitement liées qu'il n'est pas facile de les séparer. Pour cette raison nous devons traiter toutes ces catégories flexionnelles, au moins grossièrement. Nous allons donc brosser un tableau général de la flexion verbale, afin de montrer comment le temps s'insère dans ce système.

1. Nous utilisons l'expression «les formes de l'indicatif» conformément à la tradition établie. Cependant, elle ne reflète pas notre vision de la conjugaison française puisque, nous le verrons plus loin, nous considérons l'indicatif comme un grammème et non comme une classe contenant un ensemble de «temps». Il s'agit donc d'un abus de langage dans le cadre de notre modèle, que nous nous permettons d'utiliser afin d'alléger le texte pour désigner l'ensemble des formes qui expriment ce grammème.

Rappelons au lecteur qu'on trouve à l'Annexe A, à la fin de cette thèse, l'ensemble des formes verbales que nous considérons comme grammaticales et que nous décrivons ici. Insistons également sur l'importance de l'opposition grammème profond ~ grammème superficiel, telle qu'exposée au Chapitre IV (§ 1, p. 46). Enfin, il faut garder à l'esprit que notre méthodologie est basée sur une approche en deux étapes: d'abord échafauder un système flexionnel en se basant uniquement sur l'acception de base des grammèmes, puis chercher à décrire les autres acceptions de ces grammèmes ainsi que les combinaisons de grammèmes phraséologisées. Dans le présent chapitre, nous ne nous intéresserons qu'à la première étape, c'est-à-dire la construction d'un «squelette» sur lequel se base notre modèle. La description complète des grammèmes et combinaisons de grammèmes ne se fera qu'au Chapitre VIII.

1 La finitude

Les grammaires françaises ne mentionnent habituellement pas la catégorie flexionnelle de finitude, qui est généralement assimilée aux modes. Comme nous l'avons vu au chapitre précédent, la plupart des grammaires traditionnelles et contemporaines distinguent deux types de modes en français: les modes personnels (indicatif, subjonctif et impératif²), qui expriment la personne, et les modes impersonnels (infinitif, participe et gérondif), qui ne l'expriment pas. Or, nous l'avons dit dans notre critique du modèle traditionnel (Ch. V, § 1, p. 72), cette division cache à notre avis une différence plus profonde que la seule combinatoire des grammèmes. Il existe en effet une distinction importante entre les contributions de ces deux «types» de modes. Les «modes personnels» sont réputés refléter l'attitude du locuteur par rapport à l'énoncé; ils porteraient un sens d'ordre pragmatique (nous verrons plus loin nos réserves quant au sens de l'indicatif, § 3.1, p. 135). Les «modes impersonnels», quant à eux, ne portent aucun sens autre que celui de la lexie verbale. Ils servent à marquer le rôle du verbe dans l'organisation syntaxique

2. La plupart des auteurs contemporains ne considèrent plus le conditionnel comme un mode. Nous en discuterons plus loin (§ 3.4, p. 142).

de la phrase. Puisque ces deux «types de modes» forment deux classes sémantiquement bien distinctes et qu'ils s'opposent également par leur combinatoire (les uns acceptent la flexion en personne et en nombre, les autres non), nous croyons qu'il y a ici deux catégories flexionnelles plutôt qu'une seule. Nous appellerons «mode» la catégorie correspondant plus ou moins aux «modes personnels» de la grammaire traditionnelle et nous utiliserons le terme «finitude» pour désigner la catégorie correspondant à peu près aux «modes impersonnels».

Nous adoptons la définition de Mel'čuk (1994: pp. 215 *sqq.*), basée sur la théorie de la translation de Tesnière (1959), de la **finitude** comme une catégorie flexionnelle syntaxique³ indiquant si le verbe est le sommet syntaxique d'une proposition ou, s'il est un dépendant, quel rôle il remplit (celui normalement réservé à un nom, un adjectif, etc.). Selon Mel'čuk, on trouve *dans les langues en général* les cinq grammèmes «idéaux» suivants :

- **Fini**, qui marque un verbe comme sommet syntaxique d'une proposition.
- **Infinitif**, qui marque un verbe comme dépendant de type verbal.
- **Masdar**⁴, qui marque un verbe comme dépendant de type nominal.
- **Participe**, qui marque un verbe comme dépendant de type adjectival.
- **Gérondif**, qui marque un verbe comme dépendant de type adverbial.

Les grammèmes de finitude ont en principe la propriété de «déguiser» un verbe en un élément de n'importe quelle partie du discours du point de vue de son gouverneur, mais sans que ce verbe ne perde sa nature verbale du point de vue de ses dépendants (Tesnière 1959, Kahane 2001).

On ne reconnaît en français que quatre formes qui semblent clairement correspondre à ces grammèmes «idéaux»: l'infinitif, le participe présent, le participe passé et le gérondif («*en V-ant*»)⁵. Il serait possible de voir un masdar dans la

3. Une catégorie flexionnelle syntaxique est une catégorie dont les grammèmes ne servent qu'à marquer le rôle syntaxique de surface d'un lexème. Ce concept s'oppose à celui de catégorie flexionnelle sémantique, une catégorie dont les éléments expriment des sens, c'est-à-dire des éléments de la structure sémantique (voir Mel'čuk 1994: pp. 211-213).

4. Il s'agit plus ou moins du *gerund* dans la terminologie traditionnelle anglo-saxonne. Le terme vient de la grammaire arabe.

construction «(le) fait de V-er». Toutefois, celle-ci ne se comporte pas tout à fait comme les masdars véritables qu'on peut observer dans d'autres langues, comme l'anglais, où les verbes portant cette finitude se comportent encore plus comme des noms et permettent notamment d'exprimer leur premier actant par un déterminant possessif:

- (17) a. *His arriving on time surprised everyone.*
 'son arriver-MAS à l'heure surprendre-PASSÉ tout-le-monde'
 b. * *Son fait d'arriver à l'heure a surpris tout le monde.*

Il n'y a donc pas de masdar en français selon nous.

1.1 L'infinitif

Contrairement à ce que nous venons de dire, l'infinitif est généralement considéré comme «la forme nominale du verbe» en français, puisqu'il peut occuper, dans son acception de base, des fonctions qui sont typiques des groupes nominaux: sujet [*Partir plus tôt aurait été une bonne idée*], objet direct [*Il voulait partir plus tôt*], etc. Mais le fait qu'un verbe puisse remplir la même fonction qu'un groupe nominal ne veut pas forcément dire qu'il s'est «déguisé» en nom. On peut effectivement se demander si les verbes à l'infinitif qui remplissent ces rôles syntaxiques le font tout simplement en tant que verbes ou en tant que «formes nominales d'un verbe». Examinons le cas des compléments directs et indirects. Notons tout d'abord que ce ne sont pas tous les verbes qui acceptent comme complément un verbe à l'infinitif, loin de là. Grevisse (1993: § 875-877) ne dénombre que 31 verbes qui acceptent un infinitif sans préposition⁶, 138 qui acceptent un infinitif introduit par DE et 96 par À, pour un total de 265 verbes qui peuvent avoir un infinitif sans préposition ou avec DE ou À⁷. Cela contraste fortement avec la quantité de

-
5. Nous verrons un peu plus loin (§ 1.4, p. 130) que les formes de l'indicatif, du subjonctif et de l'impératif portent aussi un grammème de finitude. Nous traiterons aussi plus loin (§ 5, p. 164) des formes composées *avoir fait*, *ayant fait*, etc., que nous laissons de côté pour l'instant.
6. Il faut noter cependant qu'il ne considère pas tous les emplois. Par exemple, on ne trouve pas dans sa liste COURIR [*Elle court chercher sa clé*] ni PENSER [*Ils pensent se marier*].
7. Notons cependant qu'un certain nombre de verbes peuvent avoir un complément à l'infinitif introduit par d'autres préposition, par exemple POUR [*Ils ont voté pour ne pas réparer le toit*] ou PAR [*Ils ont fini par partir*].

verbes qui peuvent accepter un nom, notamment comme complément direct. De plus, remarquons qu'en anglais, le masdar peut occuper la place d'un nom plus facilement qu'un verbe infinitif français, par exemple :

- (18) a. *Ils ont décriminalisé la marijuana.*
 b. * *Ils ont décriminalisé fumer de la marijuana.*
 c. *They decriminalized marijuana.*
 'ils décriminaliser-PASSÉ marijuana'
 d. *They decriminalized smoking marijuana.*
 'ils décriminaliser-PASSÉ fumer-MAS marijuana'

Par ailleurs, il existe en français des verbes qui n'acceptent comme complément que d'autres verbes :

- (19) a. *Je sais nager.*
 b. * *Je sais la natation*⁸.
 (20) a. *Je pense terminer vers 17h.*
 b. * *Je pense la fin vers 17h*⁹.
 (21) a. *Il affirme savoir où se cache le voleur.*
 b. * *Il affirme sa connaissance d'où se cache le voleur*¹⁰.

Enfin, il ne faut pas voir dans *le boire*, *le savoir*, etc., la forme nominale des verbes BOIRE, SAVOIR, etc., dont il faudrait rendre compte dans notre modèle de la flexion verbale. Il s'agit clairement d'un cas de dérivation, qui d'ailleurs n'est plus productive aujourd'hui. On pourrait voir dans cette dérivation un appui à l'idée que l'infinitif soit la «forme nominale» des verbes, mais d'autres formes ont aussi tendance à se figer et à devenir des noms à part entière, soit l'indicatif (*un peigne*, (arg.) *une tire* / (québ.) *de la tire*, etc.), le participe présent (*l'équivalent*, *un correspondant*, etc.) et le participe passé (*le rendu*, *une pesée*, etc.)¹¹.

8. Les phrases du type *Je sais la douleur du solitaire* ou *Je sais l'heure qu'il est* sont construites sur des acceptions du vocable SAVOIR qui ne dénotent pas un savoir-faire, contrairement à l'acception en jeu dans les autres exemples. En d'autres termes, il ne s'agit pas de la même lexie.

9. Encore une fois, il ne faut pas confondre cette acception du verbe PENSER avec celle qu'on trouve dans, par exemple, *Ce sont eux qui ont pensé la révolution*.

10. Là non plus il ne faut pas confondre avec d'autres acceptions du même vocable [*Le Canada tente d'affirmer sa souveraineté dans l'Arctique*].

En bref, le fait qu'il y ait un certain nombre de contextes où tant les verbes à l'infinitif que les noms peuvent se retrouver ne justifie pas de considérer l'infinitif français comme indiquant qu'un verbe occupe la fonction d'un nom. Il nous semble donc que l'infinitif, dans son acception de base en français, correspond très bien à la définition de Mel'čuk, c'est-à-dire à un grammème qui indique qu'un verbe est le dépendant d'un autre verbe, sans qu'aucune translation ne soit opérée. Nous croyons qu'il y a donc un grammème **infinitif** en français, qui s'exprime par le suffixe –ER, dont les allomorphes sont *–er*, *–re*, *–oir* et *–ir*. Le choix de l'allomorphe dépend uniquement du syntactique du verbe. Ce morphème ne peut se greffer que directement à une racine verbale; il ne tolère la présence d'aucun autre suffixe flexionnel. Il peut cependant être attaché aux auxiliaires AVOIR_{acc}, ÊTRE et VENIR, dont nous parlerons un peu plus loin (nous verrons, § 5, p. 164, que ces auxiliaires ne correspondent pas vraiment à des grammèmes, ce qui veut dire que le grammème **infinitif** ne se combine à aucun autre grammème).

L'infinitif a également d'autres acceptions. Cependant, nous considérons que son acception de base est celle qui n'ajoute aucun sens au verbe. L'infinitif de consigne (typique notamment des recettes de cuisine ou encore des consignes sur les formulaires ou les affiches) exprime le sens 'il faut V-er'. Mais il est trop contraint, tant syntaxiquement que pragmatiquement, pour constituer l'acception de base du grammème **infinitif**. Quant à la construction «*de V-er*» [*Et grenouilles de se plaindre*, La Fontaine, cité dans le *Bescherelle*, § 164], il s'agit d'une expression idiomatique dont l'infinitif n'est qu'une composante qui ne porte pas en elle-même le sens que l'on reconnaît à l'ensemble.

11. Dans le cas des participes, on peut argumenter que leur conversion vers des noms ne s'est pas opérée directement, mais que les noms ont plutôt été créés à partir des adjectifs correspondants, ce qui réduirait la portée de notre argument. On peut cependant en douter dans le cas de certains noms, comme JETÉE, PESÉE, CRIÉE, RENDU, etc., pour lesquels il n'existe pas d'adjectif correspondant en français contemporain (d'après le *Petit Robert*, ils auraient été créés à partir de verbes).

1.2 Les participes

Le participe passé et le participe présent (en dehors de constructions figées comme le passé composé ou le gérondif) apparaissent tous les deux comme des dépendants de type adjectival d'un nom (ou d'un pronom)¹². Le participe passé peut en outre dépendre d'une copule [*Il semble parti*], ce qui n'est pas possible pour un participe présent [**Il semble faisant des chaises*] en français. Les deux peuvent aussi apparaître dans ce qu'on appelle traditionnellement une proposition subordonnée participiale :

(22) *Le soir tombant, le berger ramène le troupeau à la bergerie*¹³.

(23) *Le spectacle terminé, les comédiens saluent le public*¹⁴.

Remarquons que ce type de construction existe aussi avec de vrais adjectifs (quoique ce type de construction soit très contraint) :

(24) *Les feuilles enfin rouges, nous pouvions partir à la chasse au canard.*

Puisque les participes sont des modificateurs, ils ne sont pas régis par la valence syntaxique du lexème dont ils dépendent. Ils se comportent, du point de vue de leur gouverneur, comme des adjectifs qualificatifs, et c'est pour cette raison qu'ils peuvent jouer des rôles syntaxiques normalement réservés aux adjectifs.

Les participes présent et passé ont tendance à se figer et à devenir des adjectifs à part entière (*une rue passante, une couleur voyante, un homme réfléchi*, etc.)¹⁵ ou encore, comme nous l'avons déjà dit, des noms (*un passant, un voyant, un dû*, etc.). L'orthographe reflète parfois cette conversion dans le cas d'un participe présent (*la page précédant_(V) ce chapitre* → *la page précédente_(Adj)*). Par contre, il n'y a pas de différence formelle entre un participe passé et son équivalent adjectival ou nominal, sauf pour *béni_(V)* → *bénit_(Adj)*, où la différence est aussi percepti-

12. L'expression «*aller V-ant*» [*Les prix allaient augmentant*] ne constitue pas un contre-exemple puisqu'il s'agit d'une expression idiomatique. Les constructions du type *Je les ai vus buvant et chantant* sont dues à une montée du dépendant participial sur le verbe, une opération syntaxique qui existe également pour les vrais adjectifs: *Je les ai vus nus*.

13. Hugo, cité par Riegel, Pellat & Rioul (1994: p. 340).

14. Cet exemple nous vient de Riegel, Pellat & Rioul (1994: p. 343).

15. Nous empruntons ces exemples à Grevisse (1993: § 886 et § 889).

ble à l'oral (*une eau qui a été bénie* [beni], *de l'eau bénite* [benit]). Ce serait le seul exemple selon Grevisse (1993). Comme pour l'infinitif, il faut voir dans ces phénomènes des cas de dérivation, et le lien entre les participes et les formes adjectivales ou nominales qui leur correspondent est d'ordre diachronique; il ne faut donc pas en tenir compte dans notre modèle.

Outre le fait que le participe présent ne puisse être le complément d'une copule, contrairement au participe passé, les deux participes se distinguent essentiellement d'une part par l'accord — le participe présent est toujours invariable alors que le participe passé s'accorde en genre et en nombre (sauf s'il est utilisé à l'intérieur d'un passé composé avec AVOIR) — et, d'autre part, par leur diathèse¹⁶:

(25) *Les journalistes, bombardant le ministre de questions, lui font perdre son sang-froid.*

(26) *Le ministre, bombardé de questions par les journalistes, perd son sang-froid.*

Nous pourrions stipuler un grammème **participe**, commun aux deux formes, qui se combinerait à une catégorie flexionnelle, disons la voix, dont les deux grammèmes superficiels indiqueraient la diathèse. Le participe présent serait alors la composition des grammèmes **participe** et **actif**, tandis que le participe passé serait la composition de **participe** et **passif**; appelons cela «l'hypothèse des participes compositionnels». C'est en fait la position que nous avons défendue jusqu'à un stade très avancé de notre recherche. Cependant, elle est difficilement tenable. Rien, dans la morphologie des participes présent et passé, n'indique qu'ils partagent un même grammème. En soi, cela n'est pas une raison pour rejeter l'hypothèse, puisqu'il existe des signes zéros; d'ailleurs, nous avons dans notre modèle plusieurs grammèmes qui n'ont pas de signifiant propre, comme nous le verrons plus loin. Cependant, l'introduction de signes zéros doit toujours se faire avec prudence. Dans le cas des participes, nous n'avons pas trouvé d'avantage à admettre l'hypothèse des participes compositionnels. En fait, cela entraînerait beaucoup de redondance dans nos règles formelles, sans jamais permettre aucune économie.

16. Nous verrons plus loin (§ 1.2.2, p. 124) que le participe passé peut aussi exprimer des sens que ne peut avoir le participe présent.

Nous avons donc abandonné cette hypothèse pour revenir à un modèle beaucoup plus près de la tradition, où les participes sont indécomposables. Par contre, la terminologie traditionnelle est trompeuse: la différence entre le participe «passé» et le participe «présent» n'est pas temporelle. En effet, les phrases (25) et (26) ci-dessus dénotent exactement la même situation, y compris du point de vue du temps. Seule la saillance communicative des éléments de la phrase changent, et cela se manifeste par le fait qu'en (25) c'est *les journalistes* qui est le sujet de la phrase, alors qu'en (26) c'est *le ministre*. Il n'y a rien d'autre ici qu'un changement de diathèse. Par contre, c'est n'est pas toujours le cas. Les verbes intransitifs peuvent très bien être mis aux participes présent et passé sans que n'intervienne aucun changement de diathèse [*Les personnes venant pour le spectacle de magie doivent se mettre dans la file A ~ Les personnes venues pour le spectacle de magie ont été déçues*]. Nous allons donc adopter une terminologie neutre basée uniquement sur la forme. Nous appelons les grammèmes de finitude correspondant aux participes présent et passé, respectivement, **participe-ant** et **participe-é**¹⁷.

1.2.1 *Le participe présent*

Le participe présent, qui exprime le grammème superficiel **participe-ant**, ne doit pas être confondu avec les adjectifs ou les noms qui en sont dérivés. Les noms ont un comportement syntaxique suffisamment différent pour ne pas être facilement pris pour des participes présents. Par contre, le participe présent fonctionne syntaxiquement comme un adjectif (sauf qu'il ne peut pas prendre de copule). Toutefois, le participe présent n'est pas fléchi comme les adjectifs; il est invariable. Donc, s'il y a accord d'une forme en *-ant*, il ne peut s'agir que d'un adjectif (*cf.* Grevisse, 1993: § 887). On le voit bien quand l'adjectif a un sens ou une diathèse qui ne correspondent pas à ceux du participe présent à partir duquel il a été formé. Par exemple, dans la première paire de phrases ci-dessous, *croquant* est un verbe; l'accord n'est pas permis. Dans la seconde paire, c'est un adjectif; l'accord est obligatoire.

17. Nous continuons néanmoins de faire référence aux formes verbales en utilisant la terminologie traditionnelle.

- (27) a. *J'ai trouvé Laurence croquant une pomme.*
 b. **J'ai trouvé Laurence croquante une pomme.*
- (28) a. *Laurence mange une pomme croquante.*
 b. **Laurence mange une pomme croquant.*

Le sens du grammème **participe-ant** est vide. Ce grammème correspond en fait au grammème participe «pur» de Mel'čuk (1994) que nous avons mentionné au début de cette section (§ 1, p. 116). Nous n'en connaissons pas d'autre acceptation.

1.2.2 *Le participe passé*

Le participe passé est utilisé dans les formes du passé composé. Il s'agit cependant de formes figées, où le grammème superficiel **participe-é** n'est qu'une composante d'un signifiant plus complexe qui inclut aussi l'auxiliaire qui l'accompagne. Il ne faut pas considérer ces formes phraséologisées ici; nous y reviendrons plus loin (§ 5.2, p. 166 et § 6.4.2, p. 205). Ce qui nous intéresse ici est le participe passé utilisé en dehors du passé composé. Nous avons vu plus haut un exemple de participe passé exprimant un changement de diathèse, que nous répétons ici:

- (29) *Le ministre, bombardé de questions par les journalistes, perd son sang-froid.*

Cependant, le participe passé n'exprime pas toujours un tel changement de diathèse. En effet, on peut aussi l'utiliser avec des verbes intransitifs, pour lesquels le passif est pourtant impossible. Par exemple:

- (30) *Je la croyais partie pour Berlin.*

On peut le comparer au participe présent correspondant pour vérifier que la diathèse reste inchangée:

- (31) *Je l'imaginais partant pour Berlin.*

Le participe passé en (30) exprime plutôt un «résultatif», c'est-à-dire qu'il dénote l'état résultant d'un fait. Cette valeur résultative du participe passé est mentionnée dans à peu près toutes les grammaires du français que nous avons consul-

tées. Dans le cadre de la TST, Mel'čuk (1994) remarquait que le passif et le résultatif tendent à être intimement liés dans les langues en général :

Le résultat (escompté) d'une action transitive concerne plutôt l'objet que le sujet; les formes résultatives portent donc sémantiquement plus souvent sur l'objet, et pour cette raison elles sont souvent dérivées des formes passives, ou bien les formes passives en dérivent. Il en découle un lien sémantique entre la résultativité et la voix.

Mel'čuk (1994: p. 64)

On peut alors se demander quelle doit être l'interprétation première du participe passé: est-ce d'abord un passif ou un résultatif? Cette question est importante puisque, selon notre méthodologie, c'est le sens de base des grammèmes qui doit être considéré pour déterminer quelles sont les catégories flexionnelles et les grammèmes qui les composent (*cf.* Ch. IV, § 3, p. 52).

Creissels (2000) a démontré que «être V-é» peut être triplement ambigu: il peut s'agir d'un passé composé¹⁸, d'un passif, ou bien d'un résultatif. Encore une fois, ces faits sont bien connus, mais Creissels les présente de façon particulièrement claire et convaincante. Il montre d'abord que ces trois valeurs du participe passé sont relativement faciles à identifier. Le passé composé se reconnaît à la possibilité de le remplacer par une forme simple sans entraîner de différence autre qu'un changement de valeur temporelle et aspectuelle. Creissels (2000: p. 133) donne les exemples suivants:

- (32) a. *Mon ami est venu* → *Mon ami vient*
 b. *Mon ami était venu* → *Mon ami venait*

Le passif, quant à lui, peut être remplacé par un actif avec un simple changement de diathèse (on utilise ON comme sujet quand l'agent n'est pas explicitement donné dans la phrase passive). Creissels (*ib.*) propose les exemples suivants:

- (33) a. *Les coupables sont recherchés* → *On recherche les coupables*
 b. *Ils sont recherchés par la police* → *La police les recherche*

18. Nous reviendrons plus loin sur les sens du passé composé (§ 4, p. 144, § 5, p. 164 et § 6.4.2, p. 205).

Enfin, le résultatif ne se laisse pas remplacer de la même façon, comme en témoignent les exemples suivants (toujours de Creissels):

- (34) a. *Les coupables sont arrêtés depuis trois jours.*
 b. * *Les coupables arrêtent depuis trois jours.*
 c. # *On arrête les coupables depuis trois jours.*
 d. *On a arrêté les coupables il y a trois jours.*

La troisième phrase ne peut remplacer la première que si celle-ci est interprétée comme un passif (il faut alors imaginer qu'arrêter les coupables est un processus qui prend plusieurs jours), et non comme un résultatif¹⁹. Le résultatif d'un verbe transitif est toujours ambigu de cette façon, mais ce n'est pas le cas pour les verbes intransitifs, pour lesquels il n'existe pas de passif²⁰:

- (35) a. *Elle est revenue depuis deux jours.*
 b. * *On la revient depuis deux jours.*

L'impossibilité pour le participe passé d'un verbe intransitif d'exprimer la voix passive porte à croire que l'opposition participe passé ~ participe présent n'en est pas une de voix. Cependant, à ces participes passés sans interprétation passive s'opposent des participes passés sans interprétation résultative. En effet, dès qu'on exprime l'agent dans une construction passive, l'interprétation résultative devient impossible. Comparons les deux phrases ci-dessous (adaptées de Creissels 2000: p. 138). La première est ambiguë, puisqu'elle peut aussi bien dénoter un passif (la maison est en construction) qu'un résultatif (la maison est déjà construite). La seconde ne peut s'interpréter que comme un passif.

- (36) *Ma maison est construite.*
 (37) *Ma maison est construite par Constructions Beauvais Inc.*

19. La troisième phrase peut également avoir une interprétation itérative, c'est-à-dire que depuis trois jours on arrête les coupables, alors qu'avant on les laissait filer. Mais avec cette interprétation, elle ne peut pas remplacer la première phrase, peu importe le sens qu'exprime cette première phrase.

20. Ces formes sont toutefois ambiguës d'une autre manière: elles peuvent être interprétées soit comme un résultatif, soit comme un passé composé. La présence de l'adverbe DEPUIS favorise l'interprétation résultative.

La construction «*être V-é*» ne s'interprète pas de la même façon selon le type de verbe en jeu (transitif, intransitif non pronominal ou essentiellement pronominal). Le tableau suivant donne les interprétations possibles pour chaque type de verbe.

	Passif	Résultatif	Passé composé
Transitif	✓	✓	*
Intransitif	*	✓	✓ ⁱ
Pronominal	*	✓	*

Tableau XIV — Les interprétations de «*être V-é*»

i. Seulement si le verbe se conjugue avec ÊTRE.

Il ressort de ce tableau que l'interprétation résultative de «*être V-é*» semble être compatible avec une plus grande variété de verbes. Il y a cependant quelques nuances à apporter.

D'abord, cette interprétation n'est possible que si le verbe en question est télique (sinon, la construction est même agrammaticale, à moins que le verbe ne se conjugue avec ÊTRE, ce qui ne concerne qu'une minorité de verbes intransitifs non-pronominaux)²¹. La première paire de phrases ci-dessous l'illustre avec des verbes transitifs, la seconde avec des verbes intransitifs non-pronominaux, et la troisième avec des verbes pronominaux. Dans les trois paires, le verbe en *a.* est télique alors que celui en *b.* ne l'est pas.

- (38) a. *La voiture est réparée.*
 [= 'on répare la voiture' ou 'on a réparé la voiture']
- b. *La voiture est conduite.*
 [= 'on conduit la voiture', mais pas 'on a conduit la voiture']
- (39) a. *La voiture est arrêtée.*
- b. * *La voiture est accélérée.*
- (40) a. *Il est évanoui*²².
- b. * *Il est souvenu.*

21. Les pronominaux se conjuguent aussi avec ÊTRE, mais nous nous intéressons ici à la construction «*être V-é*», sans SE, c'est-à-dire à **il est souvenu*, et non *il s'est souvenu*.

22. Creissels (2000: p. 135).

Or, ce ne sont pas tous les verbes téliques qui connaissent l'emploi résultatif. Creissels (*ib.*) donne comme exemples les verbes VENIR (*il est venu* ne s'interprète que comme un passé composé) et SÉCHER (on utilise l'adjectif SEC au lieu du participe passé *séché* pour exprimer ce sens: **les chemises sont séchées, mais les chemises sont sèches*). Donc, même si le Tableau XIV indique que l'interprétation résultative de «être V-é» est compatible avec tous les types de verbes, il faut considérer que c'est en fait vrai seulement pour certains des verbes téliques.

Ensuite, il faut noter que les participes passés à valeur résultative ont tendance à se figer et à devenir des adjectifs. Par exemple, *mort* dans *il est mort* n'est pas vraiment le participe passé résultatif de MOURIR, il s'agit plutôt d'un adjectif²³. D'ailleurs, cette forme s'est également figée en tant que nom, dérivé justement de cet adjectif (c'est en tout cas la lecture du *Petit Robert*). De la même façon, *pourrie* dans *la viande est pourrie* est plutôt une forme de l'adjectif POURRI, et non du verbe POURRIR. En effet, il est clair qu'il existe un tel adjectif, comme en attestent les différents sens qu'on reconnaît dans *un été pourri, un maire pourri*, etc., pour lesquels le vocable verbal **POURRIR** n'a pas d'acception équivalente [**l'été va encore pourrir cette année, *ce maire pourrit*]. Il y a donc beaucoup de cas où il n'est pas clair si on a affaire à une construction «être V-é» ou à «être Adj».

Il reste cependant que les verbes transitifs téliques, dans cette construction, se laissent plus aisément interpréter comme résultatifs que comme passifs (*cf. La voiture est réparée*, ci-dessus, dont l'interprétation résultative semble plus naturelle). D'après le principe de l'interprétation spontanée que nous avons formulé au Chapitre IV (§ 4.1, p. 53), ce fait suggérerait que l'interprétation de base du participe passé, c'est-à-dire celle qui doit guider l'élaboration de notre modèle, serait l'interprétation résultative. Cependant, comme ce phénomène ne concerne qu'une sous-classe de verbes, il ne peut s'agir d'un critère décisif.

23. Il peut aussi s'agir d'un passé composé [*Il est mort le 8 juin*], mais nous laissons de côté cette construction pour l'instant.

Étant donné que le participe passé ne peut exprimer un sens de résultativité que pour un certain nombre de verbes téliques, alors que l'interprétation passive est compatible avec la majorité des verbes transitifs (en fait, la majorité des verbes tout court, puisque les verbes transitifs sont de loin les plus nombreux), nous concluons en nous appuyant sur le principe de la polyvalence lexicale (*cf.* Ch. IV, § 4.2, p. 57) que l'interprétation de base du participe passé dans la construction «être V-é» est celle d'un passif.

En conclusion, le grammème **participe-é**, dans son acception de base, ne diffère de **participe-ant** que par le changement de diathèse qu'il signale. Ce grammème nous intéresse dans la mesure où le participe passé se trouve impliqué dans divers phénomènes linguistiques liés au temps grammatical et à l'aspect. En effet, il est le signifiant ou une composante du signifiant de plusieurs signes :

- La voix passive, dont nous parlerons plus loin (§ 2, p. 132), s'exprime par ce grammème.
- Le passé composé, qui correspond en fait à plusieurs signes, tous pertinents pour notre étude (nous y reviendrons plus loin, § 5.2, p. 166 et § 6.4.2, p. 205), s'exprime par la combinaison d'un auxiliaire et du grammème superficiel **participe-é**. Il s'agit d'un phrasème grammématique.
- Le résultatif, dont le signifié est d'ordre aspectuel, s'exprime aussi par ce grammème. Quand le verbe est transitif, un changement de diathèse s'observe également (on observe alors la même diathèse que pour le passif).

1.3 Le gérondif

Le cas du gérondif est plus simple. Puisqu'il s'exprime toujours de façon analytique (par la préposition EN suivie d'un participe présent), il n'y a pas de grammème superficiel qui lui soit propre. Il existe néanmoins un grammème profond **gérondif** qui correspond tout à fait à celui de Mel'čuk, c'est-à-dire qu'il marque un verbe qui apparaît comme un dépendant de type adverbial d'un autre verbe. Cela peut exprimer soit la simultanéité [*Elle marchait en chantant*], soit la manière

[*Il enfonçait le bouchon en le faisant tourner*]; toutefois, nous ne sommes pas tout à fait convaincu d'avoir affaire à deux acceptions distinctes, et nous les traitons comme deux interprétations du même sens. Nous ne connaissons pas d'autre acception à ce grammème profond, qui ne se combine à aucun autre grammème. Notons par ailleurs que le gérondif se fige parfois en une locution adverbiale, par exemple ⁷EN PASSANT⁷; il ne s'agit pas là d'un phénomène flexionnel.

1.4 Les verbes finis

Nous avons épuisé les formes identifiables qui indiquent le rôle syntaxique que joue le verbe dans la phrase. Nous avons vu qu'il existe un grammème pour indiquer qu'un verbe est un dépendant de type verbal, adverbial ou adjectival. Qu'en est-il des verbes qui occupent leur position «naturelle», c'est-à-dire celle de sommet syntaxique d'une proposition? Ce ne sont habituellement que ceux à l'indicatif, au subjonctif ou à l'impératif qui peuvent occuper cette position²⁴. Or, nous l'avons dit au début de la présente section, nous ne voulons pas considérer ces formes comme faisant partie du même paradigme que l'infinitif, le participe et le gérondif, puisque d'une part elles portent des sens qui ne cadrent pas bien dans la catégorie de finitude, et d'autre part, elles expriment toujours la personne et le nombre, ce qui n'est pas compatible avec les grammèmes de finitude que nous venons de voir.

Comme il n'y a pas de morphème clairement identifiable que l'on puisse associer à la position de sommet syntaxique d'une proposition, on pourrait postuler que les verbes qui occupent cette position ne sont pas fléchis en finitude (mais le sont en mode). Cela évite de stipuler un grammème qui ne soit associé ni à un signifié²⁵, ni à une forme. On obtient alors un modèle où les verbes sont fléchis en

24. L'infinitif de consigne le peut également. Nous avons vu qu'il ne s'agissait pas de l'acception de base du grammème **infinitif**. Nous croyons en fait que cette valeur de consigne n'est pas portée par l'infinitif lui-même, mais qu'il y a ici un signe dont le signifiant est une construction telle que l'infinitif est le sommet syntaxique. Par ailleurs, nous faisons abstraction des phrases du type *Mon argent, envolé!*, où le participe semble être le sommet syntaxique; il occupe en fait la position d'un adjectif (*cf. Mon tapis, tout sale!* [on l'a sali pendant mon absence]).

mode lorsqu'ils sont le sommet syntaxique d'une proposition et en finitude lorsqu'ils apparaissent comme dépendants. Or, nous avons déjà exprimé nos réticences face à un modèle où deux catégories flexionnelles sont en concurrence et où il faut faire appel à des critères syntaxiques plutôt que morphologiques pour déterminer laquelle doit être utilisée dans tel contexte (*cf.* Ch. V, § 1, p. 72). Ajoutons que dans un tel système, le fait qu'un verbe soit fléchi en mode devient lui-même significatif, puisque cela indique que le verbe est le sommet syntaxique d'une proposition. La flexion en mode d'un verbe devient alors le signifiant d'un signe dont le signifié est le fait que le verbe est le sommet d'une proposition.

Nous considérons donc qu'il existe un grammème **fini**, dont le signifié est le fait que le verbe soit le sommet d'une proposition et dont le signifiant est nul. La combinatoire de ce grammème exige la présence d'un grammème de mode. Nous ne connaissons pas d'autre acception à ce grammème.

1.5 Synthèse

Nous avons vu que notre modèle comprend quatre grammèmes profonds de finitude. Trois sont marqués morphologiquement et un l'est syntaxiquement :

- Le grammème **fini** indique que le verbe est le sommet syntaxique d'une proposition. Il est marqué par un affixe zéro²⁶ et sa combinatoire exige la flexion en mode. Exemples : *savez*, *sachiez*, *sachez* !.
- Le grammème **infinitif** marque les verbes qui dépendent d'un autre verbe et sont régis par la valence syntaxique de ce dernier. Ce grammème est marqué formellement par le suffixe –ER (dont les allomorphes variés sont choisis selon le syntactique de la racine). Exemple : *Il veut partir*.
- Le grammème **participe-ant** marque les modifieurs verbaux qui se comportent syntaxiquement comme des verbes par rapport à leurs dépendants

25. Nous voulons dire par là qu'il n'exprime pas un élément de la structure sémantique. Cela ne l'empêche pas d'avoir néanmoins comme signifié une configuration syntaxique (bien que ce signifié soit pour le moins abstrait, il ne l'est pas plus que celui des marques d'accord, par exemple).

26. Il est possible également qu'il s'exprime de façon cumulative avec le mode et le temps. Cette question relève de la morphologie et ne nous intéresse donc pas dans cette thèse.

mais comme des adjectifs par rapport à leur gouverneur (généralement nominal). Il n'est pas compatible avec la copule, n'entraîne aucun changement de diathèse et est marqué par le suffixe –ANT. Exemple: *Toute personne se trouvant sur les lieux sera arrêtée.*

- Le grammème **participe-é** marque aussi les modifieurs verbaux qui se comportent syntaxiquement comme des verbes par rapport à leurs dépendants mais comme des adjectifs par rapport à leur gouverneur (généralement nominal). Par contre, il est compatible avec la copule, il peut entraîner un changement de diathèse et il s'exprime par le suffixe –É. Exemple: *J'habite dans une maison construite par mes grands-parents.*
- Le grammème profond **gérondif** marque les modifieurs verbaux qui se comportent syntaxiquement comme des verbes par rapport à leurs dépendants mais comme des adverbes par rapport à leur gouverneur (généralement verbal). Il s'exprime de façon analytique par la construction «*en V-ant*», c'est-à-dire en s'appuyant sur le grammème de surface **participe-ant**. Exemple: *Pierre rentre en sifflant.*

La finitude est pertinente pour l'étude du temps en français pour deux raisons. D'abord, parce que le grammème **participe-é** est utilisé dans des constructions qui servent à exprimer des sens temporels ou aspectuels. Ensuite, parce que seuls les verbes finis peuvent se combiner au temps. En effet, comme nous l'avons mentionné, les verbes finis requièrent un mode, dont un, l'indicatif, exige à son tour un grammème de temps. Nous y reviendrons plus loin (§ 3, p. 134), mais d'abord, nous discuterons brièvement la catégorie flexionnelle de la voix.

2 La voix

Dans le cadre de la TST, Mel'čuk (1994, 1997c) et Kahane (1998) définissent la **VOIX** comme une catégorie flexionnelle dont les grammèmes indiquent un changement de diathèse du lexème, sans que le sens propositionnel ne soit affecté²⁷:

(41) *Les Forces Canadiennes bombardent le pont.*

(42) *Le pont est bombardé par les Forces Canadiennes.*

La grammaire traditionnelle reconnaît en français au moins la voix active et la voix passive, auxquelles certains auteurs ajoutent la voix réfléchi (souvent appelée « pronominale »). Cependant, Mel'čuk (2001b) a démontré qu'en fait il existait en français au moins six voix (tous les exemples ci-dessous sont de lui)²⁸:

- **Actif**: *Jean a rasé Alain.*
- **Passif₁** (passif à promotion totale): *Alain a été rasé par Jean.*
- **Passif₂** (passif à promotion totale sans agent): *Une barbe de deux jours se rase sans difficulté.*
- **Passif₃** (passif à démotion partielle): *Il a été procédé par le gouvernement au licenciement des fonctionnaires inutiles.*
- **Réfléchi₁** (réfléchi direct): *Alain se rase.*
- **Réfléchi₂** (réfléchi indirect): *Alain s'achète une maison. Alain se rase la barbe.*

Tous ces grammèmes sont des grammèmes profonds. Il n'existe aucun grammème superficiel de voix en français, puisqu'il n'y a pas de signifiant morphologique propre à cette catégorie flexionnelle:

In French, the grammemes of voice appear only in the DSyntS; they do not appear in the Surface-Syntactic Structure (and of course not in the DMorphS). This happens because voice in French has only analytical (= non-morphological) expression.

Mel'čuk (2001b: p. 279)

La plupart des grammèmes de voix n'ont rien à voir avec l'expression du temps grammatical en français; ils ne sont donc pas intéressants pour notre étude. Nous renvoyons le lecteur à l'article de Mel'čuk (2001b) pour une présentation détaillée de chacun. Il nous intéresse surtout de noter que les **passif₁** et **passif₃**, s'expri-

27. La définition de Kahane (1998) distingue *voix élémentaire* et *voix composée*, mais il n'est pas nécessaire de les distinguer pour notre travail. Nos grammèmes correspondent aux voix composées de Kahane.

28. On peut voir dans *Alain s'est fait jeter dehors* et *Alain s'est vu interdire l'entrée au casino* deux autres voix profondes. Nous laissons ces constructions de côté ici.

ment par le grammème **participe-é**. Les réfléchis et le **passif₂** s'expriment quant à eux par le lexème SE²⁹, alors que la voix active n'a pas de signifiant particulier. Dans cette thèse, nous ne nous intéresserons qu'à la voix **passif₁**, que nous appellerons dorénavant «passif» tout court.

La plupart des grammairiens estiment que le marqueur de la voix passive est non pas un morphème, mais la combinaison de ÊTRE et d'un participe passé. Le problème est que la présence du verbe ÊTRE n'est pas nécessaire, comme le montrent la phrase (26), que nous reprenons en (43) ci-dessous, ainsi que (44):

(43) *Le ministre, bombardé de questions par les journalistes, perd son sang-froid.*

(44) *Nous courrions derrière des voyous poursuivis par la police.*

Abeillé & Godard (2002) ont analysé ce supposé «auxiliaire du passif» ÊTRE. Elles ont démontré qu'il a exactement les mêmes propriétés syntaxiques (et morphologiques) que la copule, notamment en ce qui a trait à la cliticisation des compléments. D'ailleurs, on peut coordonner un adjectif et un passif, comme le montre cet exemple d'Abeillé & Godard (2002 : p. 421):

(45) *Paul est aimé par tous ses collègues et prêt à les aider.*

Donc, nous croyons que la copule n'est pas inhérente aux formes passives; elle n'apparaît que parce que les participes passés se comportent syntaxiquement comme des adjectifs et ont donc besoin du support d'une copule puisqu'ils ne peuvent pas être le sommet syntaxique de la phrase.

3 Le mode

Nous avons vu lors de notre discussion de la finitude (§ 1, p. 116) qu'il existait des différences suffisantes entre les «modes impersonnels» et «modes personnels» de la grammaire traditionnelle pour les classer dans deux catégories flexionnelles différentes (les premiers relevant de la finitude, les second du mode à proprement parler). En séparant bien la finitude et le mode, nous croyons pouvoir mieux cerner

29. Voir Kayne (1977) pour la distinction entre les pronoms SE et SOI.

ce dernier. Nous n'allons cependant pas trop nous attarder sur cette catégorie, puisqu'elle n'est que secondaire dans notre recherche; nous nous contenterons de clarifier certains points liés au temps.

3.1 L'indicatif

On présente souvent l'indicatif comme étant du domaine du réel, en opposition au subjonctif, qui serait du domaine du virtuel. Nous ne connaissons pas l'origine de cette idée, mais elle perdure. On la retrouve encore, par exemple, chez Leeman-Bouix :

Le procès au subjonctif reste de l'ordre de la virtualité, du possible [...], tandis que le procès à l'indicatif est de l'ordre de la réalité, puisqu'il est inscrit dans une époque.

Leeman-Bouix (1994: p. 85)

Wagner et Pinchon (1962) attaquent cette hypothèse par des contre-exemples apparents :

La chose ainsi posée au moyen du mode indicatif peut être réelle (il a plu, il pleuvait, il pleut) ou incertaine (il pleuvra, il pleuvrait); elle peut être vraie, vraisemblable ou manifestement fausse. On voit à quel point il est inexact de définir l'indicatif comme « le mode du réel » ou « de la réalité ».

Wagner et Pinchon (1962: § 404, p. 362)

À son tour, Touratier (1996) a aussi attaqué cette idée en rappelant que le conditionnel, qui est justement utilisé pour dénoter des faits qui ne sont pas forcément du réel, est considéré comme un temps de l'indicatif par les mêmes grammairiens qui affirment que l'indicatif est du domaine de la réalité. Il note également qu'on utilise l'imparfait (de l'indicatif!) dans les subordonnées conditionnelles, qui ne dénotent justement pas des faits considérés comme faisant partie du réel [*Si tu avais mis ton foulard, tu n'aurais pas attrapé froid*].

Le problème avec ces critiques est qu'elles supposent qu'il ne peut exister qu'un seul conditionnel, et elles négligent la possibilité de formes non-compositionnelles, où l'indicatif serait privé de son sens habituel. Or, nous croyons qu'il existe plusieurs signes différents qui ont comme signifiant la forme du condition-

nel, et que certains sont justement non-compositionnels. Nous y reviendrons plus loin (§ 3.4, p. 142 et § 6.4.4, p. 213).

Mel'čuk (1994) propose une version un peu différente de la définition classique de l'indicatif. Selon lui, ce mode «exprime une assertion (= une proposition au sens logique), et pour cette raison l'énoncé à l'indicatif peut être vrai ou faux, ce qui n'est pas le cas pour les autres modes» (p. 129). Touratier (1996) s'oppose à ce type de définition basée sur la valeur de vérité puisque les interrogatives partielles à l'indicatif [*Où est-ce que tu as mis le beurre?*] ne peuvent pas être associées à une valeur de vérité. On peut cependant rejeter l'objection en invoquant le fait que c'est la construction interrogative qui place le sens de l'indicatif sous la portée de l'opérateur 'je te demande de me dire si...', ce qui a pour effet d'en masquer la valeur propositionnelle.

Toujours selon Touratier, l'indicatif n'est associé ni à un sens ni à une forme et n'est donc pas un morphème. À son avis, les verbes que l'on dit «à l'indicatif» ne portent en fait tout simplement pas de mode. Cette position n'est pas tenable dans notre cadre théorique, puisqu'elle nie le caractère obligatoire de la flexion, un postulat à la base de notre méthodologie (cf. Ch. IV, § 2, p. 51). Néanmoins, nous sommes d'accord avec lui sur un point: nous croyons aussi que l'indicatif n'est pas porteur d'un sens représentable au niveau sémantique. Nous croyons que si l'indicatif paraît ancré dans la réalité, c'est uniquement parce qu'il force l'utilisation de signes temporels qui, en situant les faits dans le temps, leur donne une apparence plus réelle; mais ce n'est pas le sens de l'indicatif.

Le grammème **indicatif** serait donc à la fois un signe non porteur de sens et un signe zéro, puisque son signifiant est nul, ce qui en fait un signe bizarre, mais un signe quand même puisqu'il a une combinatoire non triviale: **1)** quand un verbe possède plusieurs allomorphes pour son radical, il y en a un qui est réservé à l'indicatif (par opposition au radical du subjonctif), **2)** l'indicatif ne se combine (comme tous les modes) qu'aux verbes portant le grammème **fini** et **3)** les verbes à l'indicatif (et seulement ceux-là) doivent absolument être fléchis en temps.

Les signes à la fois zéro et vides illustrent bien l'importance du syntactique. Ce n'est en effet que par leur combinatoire que ces signes peuvent être observés. Mel'čuk (1997b: pp. 26-29) donne quelques exemples de tels signes en ukrainien, en russe et en allemand (il évoque également la possibilité d'analyser la forme *ont* du français comme la combinaison d'un radical vide et zéro et d'un suffixe *-ont*, qu'on retrouve par exemple dans *sont* et *font*). Il ne donne cependant pas d'exemple d'affixe flexionnel vide et zéro. En fait, sa définition des signes zéros exclut explicitement les signes morphologiques zéros sans signifié. Notre grammème **indicatif**, cependant, ne contrevient pas vraiment à cette définition. En effet, bien qu'il ne soit pas porteur d'un *sens*, il n'est pas dénué de *signifié*; seulement, son signifié n'est pas d'ordre sémantique, mais syntaxique. Le grammème **indicatif** révèle l'absence de gouverneur régissant le subjonctif. Il s'oppose dans cette position au grammème **subjonctif**, que nous allons maintenant discuter.

3.2 Le subjonctif

La concurrence du subjonctif et de l'indicatif demeure envers et contre tout l'un des trois ou quatre filons inépuisables de la linguistique française. [...] À lui seul, le rappel des méthodes, des opinions, des arguments-arguties et des polémiques prendrait un gros livre.

Wilmet (2003: p. 321)

Loin de nous l'idée d'écrire un tel livre. En fait, le subjonctif étant largement accessoire pour notre étude du temps grammatical, nous nous contenterons de lui consacrer quelques pages.

On peut dire qu'il existe parmi les grammairiens deux pôles principaux qui s'opposent au sujet du subjonctif. Il y a d'un côté ceux qui croient qu'il exprime un sens, et de l'autre ceux qui pensent au contraire que son alternance avec l'indicatif est strictement régie par le syntactique de son gouverneur.

Il est assez généralement admis parmi les défenseurs de l'hypothèse sémantique que le subjonctif «indique que le locuteur ne s'engage pas sur la réalité du fait» (Grevisse, 1993: § 864), voire qu'il présente les faits «comme une pure et simple conception de l'esprit» (Martinet, 1979: p. 120). On trouve également toutes sor-

tes de sens au subjonctif, mais bien souvent ces sens sont plutôt ceux des contextes dans lesquels s'utilisent le subjonctif, comme le notent Riegel, Pellat & Rioul :

On donne au subjonctif des valeurs qui viennent du contexte où il est employé: subjonctif de volonté (Je veux qu'il vienne), de souhait (Je souhaite qu'il vienne), de regret (Je regrette qu'il vienne), de doute (Je doute qu'il vienne); pourquoi ne pas parler aussi d'un indicatif d'opinion (Je crois qu'il viendra), d'affirmation (Je dis qu'il viendra), etc.? De fait, on attribue au subjonctif la valeur sémantique de son verbe régisseur, en confondant l'effet et la cause.

Riegel, Pellat & Rioul (1994: p. 322)

Les deux principaux arguments en faveur de l'hypothèse sémantique se trouvent résumés ici :

En premier lieu, l'emploi du subjonctif en proposition indépendante, même restreint, ne peut pas être expliqué par [la combinatoire du gouverneur]. Ensuite, il existe diverses constructions où le locuteur peut choisir entre l'indicatif et le subjonctif; le choix du mode est dans ce cas significatif: Penses-tu qu'il viendra? / qu'il vienne? - Madame de Rênal est la seule qui l'a aimé / qui l'ait aimé. La manière d'envisager le procès de la subordonnée est différente selon que le locuteur emploie l'indicatif ou le subjonctif. Le choix de ces deux modes n'est pas mécanique.

Riegel, Pellat & Rioul (1994: p. 321)

Tout d'abord, le subjonctif en proposition principale est d'un usage très restreint. En fait, il s'agit souvent de formes figées. On trouve couramment dans les grammaires SOIT [*Soit un triangle isocèle*], VIVE [*Vive le roi!*], «*AINSI SOIT-IL*», PUISSE [*Puisse le ciel nous venir en aide*] ou encore «*QUE JE SACHE*» [*Un peu d'eau gazeuse n'a jamais tué personne, que je sache*]. Pourtant, il est clair que ces formes sont lexicalisées (par exemple, elles ne peuvent pas être modifiées [**que je sache très bien*]), aussi doit-on les ignorer. Les phrases où le verbe de la principale au subjonctif exprime un ordre ou un souhait [*Qu'on fasse venir un médecin!*] ne nous aident pas non plus à déterminer le sens de base du subjonctif puisque, justement, le subjonctif n'a cette valeur d'impératif que dans cette construction bien précise, ce qui ne satisfait pas le critère de la polyvalence syntaxique énoncé plus haut (Ch. IV, § 4.2, p. 57). Même raisonnement pour les phrases du type *Moi, que je fasse une chose pareille!*.

Pour ce qui est des contextes où le subjonctif peut commuter avec un indicatif, ils ne sont intéressants que si effectivement cette alternance entraîne une différence sémantique. Il y a des lexies dont le régime laisse le choix au locuteur entre l'indicatif et le subjonctif, par exemple :

- (46) a. *Le fait que Paul soit parti chagrine Marie.*
 b. *Le fait que Paul est parti chagrine Marie.*

Certains verront dans les alternances de ce type une nuance de sens. Leeman-Bouix (1994), à qui nous empruntons l'exemple, croit que dans la première phrase le fait est présenté comme possible, alors que dans la seconde le locuteur est certain de ce qu'il affirme (un point de vue assez répandu, notamment chez les héritiers de Guillaume). Il nous semble plutôt que ces deux phrases présentent le départ de Paul comme un fait avéré. Si le fait dénoté par un verbe à l'indicatif peut sembler plus réel, comme nous l'avons dit plus haut (§ 3.1, p. 135), c'est à notre avis uniquement parce que ce mode se combine aux grammèmes de temps, ce qui permet de situer temporellement les faits, alors que ce n'est pas possible au subjonctif³⁰. Les faits étant ainsi ancrés dans le temps, ils paraissent mieux ancrés dans la réalité. Mais ce n'est pas le subjonctif qui porte un sens 'possible' ou 'virtuel'.

Ensuite, certaines alternances en apparence libres entre l'indicatif et le subjonctif sont en fait dues à des gouverneurs homonymiques n'ayant pas le même régime. Par exemple, le vocable **EXPLIQUER** dans les phrases suivantes n'a pas le même sens. Il s'agit donc de deux lexèmes distincts, et chacun a son régime propre: le premier commande l'utilisation de l'indicatif, alors que le deuxième régit le subjonctif.

- (47) a. *Paul expliquait qu'il n'avait pas pu arriver plus tôt.*
 b. *Voilà qui explique que vous n'ayez pas pu arriver plus tôt.*

Menanteau (1986) a fait une expérience pour tester l'hypothèse selon laquelle l'alternance libre indicatif~subjonctif serait significative, en soumettant des paires

30. L'auxiliaire AVOIR du subjonctif passé est d'ordre aspectuel, comme nous le verrons plus loin (§ 5, p. 164).

de phrases ne différant que par cette alternance à des locuteurs natifs. Ses résultats semblent confirmer notre intuition :

Tirée à 600 exemplaires, cette enquête a été diffusée dans des milieux socio-professionnels et socioculturels diversifiés et l'on constate que les différences de sens ne sont repérées que dans 28 % des cas. Ainsi, dans près de trois cas sur quatre, lorsque le sujet parlant a le choix entre le subjonctif et un autre «mode», ce choix ne correspond pas pour lui à une différence de sens. Le subjonctif est alors simple variante libre; son usage relève davantage de la sociolinguistique: qui l'emploie marque par là son appartenance (réelle ou fantasmée) à tel ou tel groupe plus prestigieux.

Menanteau (1986: pp. 72-73)

Malheureusement, il ne détaille pas les résultats de son enquête dans cet article (et nous ne les avons pas trouvés ailleurs). Encore plus regrettable, l'expression «28 % des cas» est ambiguë: il n'est pas tout à fait clair pour nous si cela signifie que seulement 28 % des répondants ont perçu des différences sémantiques dans les paires de phrases soumises, ou si les répondants en général n'ont perçu des différences sémantiques que dans 28 % de ces paires (nous croyons que c'est cette dernière interprétation qui est la bonne). Quoi qu'il en soit, cela nous indique que dans la plupart des cas, l'alternance libre indicatif~subjonctif n'est pas significative; elle est plutôt de l'ordre du registre. Il reste néanmoins des alternances du type suivant (nous empruntons cet exemple à Menanteau):

- (48) a. *Je cherche quelqu'un qui soit parisien.*
 b. *Je cherche quelqu'un qui est parisien.*

Ici, la nuance sémantique est assez claire. La première phrase indique que je cherche n'importe qui, du moment que cette personne soit de Paris (ici, le conditionnel pourrait aussi être utilisé); j'exprime que je n'ai pas de référent précis en tête. La seconde est ambiguë: elle peut avoir le même sens que la première ou (peut-être plus naturellement) elle peut signifier que je cherche une personne en particulier, et que cette personne est parisienne. Il n'y a pas de doute qu'il y a une unité significative en jeu dans cette opposition entre indicatif et subjonctif. Cependant, le subjonctif ne peut avoir ce sens que dans un contexte bien précis: il doit se trouver dans une subordonnée restrictive dont le gouverneur est le complément (direct ou indirect) d'un verbe (ce complément peut éventuellement avoir été

«déplacé», dans une dislocation par exemple). Comme pour le subjonctif à valeur d'impératif, ce subjonctif à valeur «référentielle» est d'un usage trop restreint pour servir d'acception de base.

En conclusion, nous croyons que dans son acception de base, le subjonctif est vide de sens. L'alternance subjonctif~indicatif est généralement régie par la lexie dont dépend le verbe. La construction dans laquelle apparaît le verbe peut également admettre ce mode [*Je pense que tu peux venir ~ Je ne pense pas que tu puisses venir*]. Utilisé en proposition principale, le subjonctif exprime un ordre ou un souhait; il sert à compléter le paradigme défectif de l'impératif. Utilisé dans une subordonnée restrictive, il indique que le référent du nom dont dépend la subordonnée n'est pas connu.

3.3 L'impératif

Les formes de l'impératif coïncident presque toujours avec celles de l'indicatif présent ou du subjonctif. Il n'y a que deux exceptions, *sachons* et *sachez*, qui sont construites avec le radical du subjonctif et le suffixe de l'indicatif présent. Sans ces deux formes mixtes, il serait possible de modéliser l'impératif directement dans la transition entre les niveaux sémantique et syntaxique en faisant correspondre au sens de l'impératif le grammème **subjonctif** ou la combinaison **indicatif** ⊕ ***présent**³¹, selon le verbe. On obtiendrait alors un modèle où il n'y aurait pas de grammème réservé à l'impératif. Cependant, les formes *sachons* et *sachez* ne peuvent pas être traitées de cette façon: il faut avoir au niveau syntaxique un grammème **impératif** qui sera mis en correspondance au niveau morphologique avec les morphèmes du subjonctif ou de l'indicatif et du présent, selon le verbe, et qui forcera la sélection du bon radical (le radical, dans ce cas, ne peut pas être imposé simplement par le morphème de l'indicatif ou du subjonctif). De toute façon, le grammème **impératif** a clairement un impact sur la syntaxe de la phrase puisqu'il

31. Nous verrons plus loin (§ 6.4, p. 202) que ce que nous appelons ici «présent» correspond en fait à deux grammèmes (**simultané** et **non-décalé**).

force l'absence de sujet. On ne peut donc pas faire l'économie de ce grammème dans notre modèle.

Nous ne connaissons qu'une seule acception au grammème **impératif**; il sert à exprimer une «requête», au sens large :

(49) *Passe-moi le sel.*

(50) *Essayez en tournant de l'autre côté, peut-être.*

Ce grammème ne se combine qu'à la deuxième personne du singulier et du pluriel ainsi qu'à la première personne du pluriel.

3.4 Le conditionnel

Le conditionnel appelle quelques commentaires. Longtemps considéré comme un mode, il est aujourd'hui rangé parmi les temps de l'indicatif par la plupart des grammaires contemporaines. Entre autres, Grevisse (1993), Leeman-Bouix (1994) ou encore Riegel, Pellat et Rioul (1994) soutiennent que les formes du conditionnel expriment des sens tant temporels [*Ils nous ont dit qu'ils finiraient le lendemain*] que modaux [*J'en prendrais bien un peu*], tout comme d'autres formes de l'indicatif, tel le futur simple [*Il viendra demain* (sens temporel) ~ *Il se sera sans doute perdu* (sens modal)], et en concluent qu'il s'agit donc d'une forme de l'indicatif. Par ailleurs, Le Goffic (1997) estime que le conditionnel a d'abord une valeur d'affirmation sous condition qui le rangerait parmi les formes exprimant un jugement, ce qu'il considère être le propre de l'indicatif.

Nous sommes également d'avis qu'il faut ranger le conditionnel parmi les formes de l'indicatif, mais pas tout à fait pour les mêmes raisons. Bien qu'il est vrai que le conditionnel, tout comme certaines formes de l'indicatif, peut porter aussi bien des sens temporels que modaux, il reste qu'il est fréquemment utilisé comme marqueur de sens typiquement modaux (par exemple, le doute ou la réserve [*Le suspect aurait fait feu en direction des policiers*] — très commun dans les textes journalistiques — ou encore l'atténuation [*Je prendrais une frite s'il vous plaît*]), alors que les autres formes de l'indicatif ne portent pas aussi fréquemment de sens

modaux. Qui plus est, comme le note Gosselin (1996: p. 223), le conditionnel ne porte un sens temporel que dans des propositions subordonnées [*Ils ont dit qu'ils viendraient*]. Lorsqu'un verbe principal est au conditionnel, il exprime presque toujours un sens modal, et non temporel [*Ils viendraient* 'il semble qu'ils viennent']. Ce n'est évidemment pas le cas des autres formes synthétiques de l'indicatif, qui peuvent porter un sens temporel peu importe leur position dans la structure syntaxique.

Selon les critères de l'interprétation spontanée et du sommet syntaxique (*cf.* Ch. IV, § 4, p. 53), ces deux faits nous porteraient à croire que les formes du conditionnel expriment plus un sens modal que temporel et que nous avons affaire à un mode qui peut à l'occasion exprimer des sens temporels, et non à un temps qui peut exprimer des sens modaux (comme le fait le futur simple, par exemple). Toutefois, dès qu'on prend en considération les formes analytiques de l'indicatif (que nous présenterons dans les sections qui suivent), on constate que le conditionnel à valeur temporelle n'est pas le seul à ne pouvoir apparaître que dans des propositions subordonnées: c'est aussi le cas notamment des formes construites avec un auxiliaire AVOIR à l'imparfait (plus-que-parfait et plus-que-parfait surcomposé). Cette restriction peut s'expliquer par le caractère «anaphorique» du sens dénoté par ces formes (nous y reviendrons plus loin, § 6, p. 177). Quant à la fréquence relativement élevée des occurrences du conditionnel à valeur modale par rapport aux occurrences à valeur temporelle, elle s'explique aisément par l'utilité relative de ces sens dans les situations réelles de communication. Il est effectivement plausible qu'il soit plus souvent nécessaire d'exprimer la politesse, le doute ou la réserve, par exemple, que de situer un fait dans le temps d'une façon aussi complexe que le fait le conditionnel. Le sens seul ne permet donc pas de décider si le sens de base du conditionnel est d'ordre temporel ou modal, c'est-à-dire s'il faut ranger le conditionnel parmi les temps ou les modes.

De toute façon, le conditionnel est un signe complexe qui peut s'analyser en la combinaison du futur simple et de l'imparfait. Il est encore trop tôt pour démontrer cette affirmation mais nous verrons plus loin (§ 6, p. 177) que dans sa valeur tem-

porelle ('futur par rapport au passé'), le conditionnel est strictement compositionnel. Dans ses emplois modaux, le conditionnel est un phrasème grammématique. Il en découle qu'il ne s'agit pas d'un grammème superficiel, mais de la combinaison de deux grammèmes superficiels; aussi la question de savoir s'il faut le ranger parmi les modes ou les temps n'a-t-elle aucun sens.

3.5 Synthèse

Nous avons donc dans notre modèle une catégorie flexionnelle de mode contenant trois grammèmes, qui s'expriment tous par un suffixe:

- Le grammème **indicatif** est sélectionné par défaut. Son signifié n'est pas sémantique et son signifiant est un suffixe zéro. Les verbes à l'indicatif doivent porter des grammèmes temporels.
- Le grammème **subjonctif**, dans son acception de base, est généralement imposé par le régime d'une lexie. Il s'exprime par le suffixe $-I-$ (dont les allomorphes sont $-i-$ aux première et deuxième personnes du pluriel et $-\emptyset-$ aux autres personnes) et sélectionne un radical particulier pour certains verbes.
- Le grammème **impératif** exprime une «requête» au sens large. Il ne s'exprime pas par un morphème particulier, mais par les suffixes $-\emptyset-$ de l'indicatif ou $-I-$ du subjonctif, selon le verbe. Il ne se combine qu'aux première et deuxième personnes du pluriel ainsi qu'à la deuxième personne du singulier.

4 Les formes surcomposées

Nous quittons maintenant le domaine des grammèmes syntaxiques puisque les formes qu'il nous reste à disséquer sont celles qui expriment des sens d'ordre temporel ou aspectuel. Ce sont elles qui nous intéressent réellement dans le cadre de cette thèse. Nous allons aborder leur étude en commençant, pour ainsi dire, par le dessert, c'est-à-dire par les formes dites «surcomposées». Il peut sembler étrange d'aborder le problème sous cet angle, puisque ces formes sont finalement assez

marginales et ne sont habituellement mentionnées qu'en dernier lieu dans les grammaires, voire pas du tout, mais nous verrons qu'elles fournissent de précieux indices sur la nature de l'auxiliaire AVOIR et sur l'articulation des catégories de mode, de temps et d'aspect.

4.1 Définition et acceptabilité des formes surcomposées

Nous entendons par «**forme surcomposée**» toute forme verbale contenant soit deux auxiliaires AVOIR, soit un auxiliaire AVOIR et un auxiliaire ÊTRE (qui n'est pas la copule d'une forme passive). Par exemple : *il a eu fini, il a été parti, il s'était eu habillé*, etc. Ces formes sont mentionnées dans la plupart des ouvrages (on en parlait déjà dans les grammaires au XVI^e siècle) mais en général peu discutées sous prétexte qu'elles sont rares.

Ces formes sont par ailleurs souvent stigmatisées en raison de leur forte association à la langue orale ou encore à un parler «rural». Pourtant, deux études systématiques et relativement récentes de Jolivet (1984) et de Hill (1984) suggèrent que les formes surcomposées sont plus utilisées et acceptées dans les groupes sociaux-économiques les plus élevés et par les locuteurs les plus instruits ou les plus vieux. Mais une étude plus récente de Carruthers (1999) indique qu'en fait la réalité est un peu plus complexe, et que les résultats dépendent beaucoup du type de test utilisé (questionnaire écrit, questionnaire oral ou entrevue dirigée). Dans le discours libre, les locuteurs ont tendance à éviter les constructions propices à l'apparition de surcomposés, alors que lorsqu'on les soumet à des questionnaires et qu'on les force à utiliser ces constructions, ils sont d'autant plus susceptibles d'utiliser ces formes qu'ils sont instruits.

Tous les auteurs ne s'entendent pas sur la liste des formes surcomposées grammaticalement correctes; il est donc nécessaire de régler cette question avant de poursuivre. En appliquant mécaniquement la surcomposition à toutes les formes composées, on obtient l'ensemble des formes surcomposées théoriquement possibles (nous faisons ici abstraction des formes avec un auxiliaire ÊTRE):

Forme simple	Forme composée	Forme surcomposée
<i>fait</i>	<i>a fait</i>	<i>a eu fait</i>
<i>faisait</i>	<i>avait fait</i>	<i>avait eu fait</i>
<i>ferait</i>	<i>aurait fait</i>	? <i>aurait eu fait</i>
<i>fera</i>	<i>aura fait</i>	? <i>aura eu fait</i>
<i>fit</i>	<i>eut fait</i>	* <i>eut eu fait</i>
<i>fasse</i>	<i>ait fait</i>	* <i>ait eu fait</i>
<i>fît</i>	<i>eût fait</i>	? <i>eût eu fait</i>
<i>faisons !</i>	<i>ayons fait !</i>	* <i>ayons eu fait !</i>
<i>faire</i>	<i>avoir fait</i>	? <i>avoir eu fait</i>
<i>faisant</i>	<i>ayant fait</i>	? <i>ayant eu fait</i>
<i>fait</i>	* <i>eu fait</i>	* <i>eu eu fait</i>

Tableau XV — Les formes surcomposées théoriquement possibles

De ces onze formes surcomposées hypothétiques, nous n'en considérons que deux comme clairement grammaticales (les deux premières). Deux autres, le «conditionnel surcomposé» et le «futur antérieur surcomposé», nous font hésiter en raison de leur usage limité, mais ne nous semblent pas agrammaticales.

La forme hypothétique **eu eu su*, construite sur une forme composée elle-même agrammaticale (**eu su*), ne présente aucun intérêt. Pour toutes les autres formes cependant, on trouve dans la littérature scientifique des exemples plus ou moins convaincants qu'il convient de passer en revue afin de justifier nos jugements de grammaticalité.

4.1.1 Le passé surcomposé

On appelle généralement ainsi les formes surcomposées construites sur la base du passé composé. Le passé surcomposé fait consensus sur au moins deux points. D'abord, bien que rare, il est de loin la forme surcomposée la plus fréquente (des 14 formes surcomposées relevées par Carruthers (1996) dans son corpus, toutes étaient des passés surcomposés). Ensuite, on lui reconnaît un sens de base qui com-

bine l'aspect accompli et l'antériorité (nous reviendrons un peu plus loin sur la question, § 4.5, p. 158), comme dans la phrase suivante :

- (51) *Quand nous avons eu fini de goûter, j'ai fait goûter Noël.*
Duras, citée par Grevisse (1993 : § 788)

Cette construction est sémantiquement équivalente à un passé antérieur :

- (52) *Quand nous eûmes fini de goûter, je fis goûter Noël.*

Elle n'est toutefois pas équivalente à un passé composé. En effet, il y a une subtile nuance entre cette phrase et la phrase ci-dessous :

- (53) *Quand nous avons fini de goûter, j'ai fait goûter Noël.*

En (53), le passé composé *avons fini* renvoie directement au moment où la collation des uns s'est terminée (moment auquel Noël a pu avoir la sienne), alors qu'en (51) on fait référence non pas à cet instant, mais à la phase qui le suit. La différence entre *j'ai eu fait* et *j'ai fait* est la même qu'entre *j'ai fait* et *je fais*.

Pourtant, bien qu'il existe clairement une différence sémantique entre le passé surcomposé et le passé composé, les locuteurs semblent avoir tendance à utiliser les deux formes indistinctement lorsqu'on les soumet à des tests linguistiques visant à provoquer l'usage du surcomposé. En utilisant des techniques différentes, Hill (1984) et Carruthers (1999) ont obtenu des résultats très similaires quant au taux d'utilisation effective du passé surcomposé. Dans les contextes où on aurait attendu cette forme, moins du quart (22,1 % selon Hill, 23 % d'après Carruthers) des répondants l'ont effectivement utilisée, contre 33,8 % (d'après Hill) qui ont utilisé un passé composé³². Pourquoi? Hill fournit une explication du phénomène: cela dépend tout simplement de la façon dont les répondants ont interprété la situation qu'on leur demandait d'exprimer. Par exemple, dans ce test, les répondants devaient compléter la phrase *Lorsqu'il (finir) son travail, il est rentré chez lui*. D'après Hill, ceux qui avaient en tête le fait de terminer le travail auraient utilisé le passé composé, et ceux qui avaient en tête l'état résultant d'avoir fini le travail auraient utilisé le passé surcomposé (ou le passé antérieur). Il faut cependant consi-

32. Autant de locuteurs ont utilisé le passé antérieur dans l'étude de Hill.

dérer aussi la possibilité que s'opère une neutralisation de ces formes chez certains locuteurs, qui peuvent par exemple utiliser les phrases (51) et (53) ci-dessus indistinctement pour décrire la même situation. L'étude de Carruthers (1996) montre très clairement qu'il y a de bonnes raisons de le croire: «The data suggest that while the *surcomposé* is used only in a minority of cases (23%), it is the *passé composé* which is used in most subordinate temporal clauses marking anteriority in relation to a main clause *passé composé* (72.1%)» (p. 188).

Fait intéressant à noter, l'étude de Hill (1984) suggère que plus les répondants sont âgés, plus ils ont tendance à utiliser le passé surcomposé. Pour la phrase *Lorsqu'il (finir) son travail, il est rentré chez lui*, seulement 14,3 % des moins de 20 ans ont utilisé le passé surcomposé, contre 22,2 % des répondants entre 21 et 50 ans, et 38,5 % des plus de 50 ans (il faut noter cependant que l'échantillon ne comptait que 68 répondants au total). L'auteur y voit une régression du surcomposé, mais d'autres explications sont à considérer. Il est possible en effet que cette forme soit acquise tard, peut-être à cause de sa rareté et de la subtilité de son sens. Il est également possible que ces différences soient dues à une plus grande confiance des personnes âgées en leurs capacités linguistiques, ce qui les ferait moins hésiter face à une forme rare et stigmatisée. Enfin, il faut considérer le fait que les personnes nées avant 1934 ont en moyenne été beaucoup moins scolarisées et sont peut-être moins conscientes d'utiliser une forme stigmatisée. Il s'agit là d'hypothèse que nous ne sommes pas en mesure de vérifier et qui dépassent le cadre de notre thèse, mais il nous semble en tout cas précipité d'annoncer le déclin du passé surcomposé.

Toutes les études dont nous avons connaissance sur le sujet ont montré l'importance du contexte syntaxique pour les formes surcomposées. Premièrement, le passé surcomposé, dans le sens que nous avons considéré jusqu'ici, ne se trouve que dans des propositions temporelles subordonnées (qui dépendent généralement d'une proposition principale au passé composé). Il a également fortement tendance à apparaître simultanément avec un marqueur explicite d'antériorité du type *quand, dès que, une fois que*, etc. (cf. Carruthers 1996). On remarque aussi que ce

sont souvent les verbes FINIR, TERMINER, etc. qui portent cette forme, mais il ne s'agit pas là d'une contrainte.

On trouve également dans presque tous les travaux sur les surcomposés des exemples du type suivant (on trouve aussi des exemples avec TÔT ou BIENTÔT au lieu de VITE):

- (54) *Ce petit vin nouveau [...] a eu vite grisé tous ces buveurs de bière.*
Daudet, cité par Grevisse (1993: § 788)

Ici, ce qui semble être un passé surcomposé se trouve dans une proposition indépendante. Mais nous croyons qu'il y a là une locution «*AVOIR VITE*» (et qu'il existe aussi les locutions «*AVOIR TÔT*» et «*AVOIR BIENTÔT*»). D'une part, ces constructions ne se laissent pas modifier librement. On ne peut pas, par exemple, remplacer VITE, TÔT ou BIENTÔT par d'autres adverbes comme RAPIDEMENT, BRUSQUEMENT, SOUDAINEMENT, ENCORE, etc.³³ On ne peut pas non plus l'utiliser sans auxiliaire. D'autre part, on peut voir dans la phrase suivante, par exemple, que le sens de cette locution n'est pas tout à fait compositionnel:

- (55) *Les « élites » off shore, en particulier en France, ont tôt fait de qualifier de « nationalisme » l'attachement des peuples à leur langue, alors que c'est parfois tout ce qui leur reste pour « faire société » et s'inscrire dans une histoire partagée.*

Manière de voir, numéro 97 (février/mars 2008), p. 5

Ici, la locution est utilisée sans deuxième auxiliaire AVOIR, ce qui simplifie un peu les choses. On voit bien que *ont tôt fait* n'est pas un banal passé composé, qu'il s'agit là d'un présent (notons le présent *c'est* dans la deuxième partie de la phrase). On peut en effet paraphraser la construction par un verbe au présent, par exemple S'EMPRESSER [*Les élites s'empressent de qualifier...*]. Ainsi, la construction qu'on trouve en (54) n'est rien d'autre que le passé composé de la locution «*AVOIR VITE*», et non pas un passé surcomposé de FAIRE.

33. Dans le cas des adverbes en *-ment*, il est possible que ce soit dû au fait qu'ils vont de préférence après le verbe. Il n'est toutefois pas impossible de les mettre avant le participe passé dans une forme composée [*Ils ont rapidement remonté la pente*].

Enfin, on mentionne dans la littérature sur le sujet d'autres sens du passé surcomposé, mais il s'agit d'usages limités à quelques variétés du français, et ils doivent donc être exclus de notre étude. Nous insistons plus loin sur ce point (§ 4.3, p. 156).

4.1.2 *Le plus-que-parfait surcomposé*

La seule autre forme surcomposée qui soit relativement bien représentée dans les corpus est le «plus-que-parfait surcomposé», construit à partir du plus-que-parfait. Cette forme est au plus-que-parfait et à l'imparfait ce que le passé surcomposé est au passé composé et au présent, tant dans le sens que dans la forme. Ses conditions d'emploi sont les mêmes que pour le passé surcomposé, si ce n'est que le verbe dont dépend la proposition temporelle où se trouve cette forme est normalement au plus-que-parfait (plutôt qu'au passé composé):

(56) *Quand il avait eu rassemblé les plus effrontés de chaque métier, il leur avait dit: régnons ensemble.*

Stendhal, cité par Grevisse (1993: § 788)

Comme pour le passé surcomposé, on a tendance à confondre le plus-que-parfait de «AVOIR VITE» avec un plus-que-parfait surcomposé:

(57) *Ah! L'idiote avait eu vite fait de se couler! Il n'avait pas fallu deux mois pour que le fils bien-aimé revînt dormir dans son petit lit de collégien.*

Mauriac, cité par Cornu (1953: p. 113)

Ce que nous avons dit plus vaut ici. Par ailleurs, il existe des emplois dialectaux de cette formes, que nous ignorons dans le cadre de cette thèse. Le lecteur intéressé par le sujet pourra consulter en particulier Cornu (1953).

4.1.3 *Le conditionnel surcomposé*

En surcomposant le conditionnel passé, on obtient le conditionnel surcomposé. D'après une étude de Jolivet (1984) sur l'acceptabilité des surcomposés, ces formes ne seraient pas particulièrement acceptés, bien au contraire, sauf en proposition relative. Pourtant, le *Dictionnaire de l'Académie française* (1878) donne un exemple d'une telle forme en proposition indépendante sous l'entrée **AVOIR**:

(58) *Sans lui, j'aurais eu dîné de meilleure heure.*

Académie française (1878)

La phrase ne nous semble pas particulièrement mauvaise, quoique peut-être peu naturelle. On peut en fournir un vague équivalent dans un style plus familier:

(59) (fam) *J'aurais eu fini bien avant si tu m'avais pas dérangé.*

La nuance par rapport au conditionnel passé (*j'aurais fini...*) est la même que pour les autres formes surcomposées déjà mentionnées: on fait référence ici non pas au moment où le locuteur termine son action, mais à la phase qui suit ce moment.

Pour ce qui est de l'usage de cette forme en proposition relative, Damourette et Pichon en citent un exemple qui nous semble tout à fait correct, où on voit encore mieux la nuance sémantique entre le conditionnel passé et le conditionnel surcomposé:

(60) *Lorsqu'il aurait eu dit: «Le roi de France et trois cent mille citoyens furent égorgés, fusillés, noyés [...]». Quels mots aurait-il mis au-dessous de pareilles choses?*

Hugo, cité par Damourette & Pichon (1911–1950: § 1862)

Remarquons que le suffixe que porte le premier auxiliaire d'un conditionnel surcomposé ne peut exprimer qu'un sens modal. Il ne s'agit pas d'un «futur du passé» [**Il a promis qu'il aurait eu fini avant 15h*]. Cela contraste avec le conditionnel passé, où le même suffixe peut très bien exprimer un sens temporel qui s'ajoute au sens aspectuel de l'auxiliaire [*Il a promis qu'il aurait fini avant 15h*].

Enfin, notons que la locution 「AVOIR VITE」 peut également se trouver au conditionnel passé, et qu'on peut alors la confondre avec le conditionnel surcomposé de FAIRE:

(61) *En cas d'alerte, chacun aurait eu vite fait de retrouver son bien.*

Bazin, cité par Cornu (1953: p. 132)

4.1.4 *Le futur antérieur surcomposé*

Plus rare encore est la forme surcomposée construite à partir du futur antérieur. Jolivet (1984) note que toutes les formes surcomposées à valeur de futur [*il aura eu fait*] ou hypothétique [*s'il avait eu fait*] sont «sévèrement jugées», dans le cadre de son étude sur l'acceptabilité des surcomposés. Son article n'insiste toutefois pas sur ce point et se concentre plutôt sur les trois autres formes dont nous venons de parler. On trouve cependant dans la littérature l'exemple suivant, qui ne nous semble pas clairement agrammatical :

- (62) *On pense que M. Tardieu en aura eu fini hier soir avec les résistances du Dr Schacht, il aura pris le train de 20 heures pour être à 6h30 à Paris, tenir à 10 le conseil des ministres.*

Maurras, cité par Damourette et Pichon (1911–1950: § 1859)

Dans cette phrase, la forme surcomposée ne peut pas être remplacée par un futur antérieur sans que n'en soit altéré le sens. Par contre, il faut noter que ce futur antérieur surcomposé ne situe pas le fait dénoté comme postérieur à quoi que ce soit. Comme pour le conditionnel surcomposé, le suffixe du futur qu'on trouve sur le premier auxiliaire du futur antérieur surcomposé ne peut avoir qu'un sens modal ; il sert à exprimer une conjecture. C'est en fait la combinaison du futur de supposition (*cf.* § 6.4.4, p. 213) et de l'accompli (*cf.* § 5.2, p. 166).

4.1.5 *Les autres «formes surcomposées»*

Les surcomposés que nous venons de voir sont à notre avis les seuls qui existent en français standard contemporain. Les autres formes mentionnées dans les travaux sur le sujet, soit relèvent de dialectes que nous ne considérons pas dans notre travail, soit nous paraissent douteuses, voire carrément agrammaticales. Malheureusement, le passé surcomposé, le plus-que-parfait surcomposé et le conditionnel surcomposé ont monopolisé l'attention des rares études systématiques qui ont été faites sur l'acceptabilité des surcomposées, si bien que nous n'avons pour nous guider que peu de données sur les autres formes surcomposées.

Le soi-disant passé antérieur surcomposé (**eut eu fait*) est, comme le note Grevisse (1993: § 788) «inconnu de la langue parlée, très rare dans l'écrit». La raison

en est tout simplement que cette forme n'est pas grammatical à notre avis. Considérons par exemple la citation suivante :

- (63) *Son fils était resté à terre pour fermer la barrière. Quand il eut manœuvré et que la voiture l'eût [sic] eu franchise³⁴, le petit courut pour grimper auprès de son père.*

Vialar, cité par Grevisse (1953 : § 788)

Nous ne voyons pas ce que ce double auxiliaire apporte de plus qu'un simple auxiliaire. Cornu (1953) est également d'avis que le passé antérieur surcomposé est incorrect, mais il croit qu'il exprime quand même une nuance qui lui est propre, mais «dont on ne voit pas nettement la nécessité» (*ib.*, p. 126).

Le prétendu subjonctif passé surcomposé (*?ait eu fait*) ne nous paraît pas non plus apporter quelque nuance que ne saurait exprimer un subjonctif passé :

- (64) *Je me serais ennuyée à mourir avant qu'il n'ait eu fini, si je n'avais pas été installée près de cette fenêtre-là.*

Paratte, cité par Grevisse (1953 : § 788)

Quant au subjonctif plus-que-parfait surcomposé (**eût eu fait*), encore une fois inconnu de la langue parlée et rare à l'écrit selon Grevisse (*ib.*), il est assez difficile d'en trouver un exemple convaincant. On trouve dans *Le bon usage* les exemples suivants :

- (65) *De par la rage de sa passion Jacques eût eu acquis des boutons sur la face s'il n'eût eu Martine pour s'exercer.*

Queneau, cité par Grevisse (1993 : § 788)

- (66) *Après que j'eusse eu fini nous fûmes invités [...] à une soirée*

Verlaine, cité par Grevisse (1993 : § 788)

Il note que l'exemple de Queneau est une plaisanterie, et que celui de Verlaine est «surprenant à tous égards» (en effet, la concordance d'un subjonctif plus-que-parfait surcomposé avec un passé simple est plutôt douteuse).

34. Grevisse interprète cette forme comme un passé antérieur surcomposé, bien que *eût* soit une forme de l'imparfait du subjonctif. L'auteur a-t-il mis un accent circonflexe où il n'en fallait pas, ou voulait-il vraiment avoir ici un subjonctif plus-que-parfait surcomposé? Quoi qu'il en soit, les deux formes nous semblent incorrectes.

Les verbes conjugués avec ÊTRE peuvent porter à confusion, puisque leur participe passé a bien souvent un adjectif homonyme. Dans les exemples suivants, nous croyons qu'on a affaire à des adjectifs utilisés avec le plus-que-parfait du subjonctif de la copule ÊTRE, et non à des formes surcomposées :

(67) *Gustin se soumettait à toutes mes volontés, comme s'il eût été né pour m'obéir.*

Quinet, cité par Grevisse (1993 : § 788)

(68) *Une heure plus tard, il eût été parti pour l'Amérique.*

Balzac, cité par Cornu (1953 : p. 143)

Enfin, le participe surcomposé et l'infinitif passé surcomposé ne sont pas de meilleurs candidats :

(69) *Ayant eu terminé son travail avant midi, il a pu avoir son train ordinaire.*

Grevisse (1993 : § 788)

(70) *Le plombier est parti sans avoir eu achevé son travail.*

Grevisse (1993 : § 788)

Dans ces deux exemples, les formes surcomposées se laissent remplacer par des formes composées sans que le sens de la phrase ne soit en rien altéré. À ce propos, citons Cornu, qui avait bien cerné le problème :

Si je veux faire entendre qu'un tel a coutume d'écrire un résumé des ouvrages qu'il lit, une fois sa lecture achevée, j'indiquerai l'antériorité de l'achèvement de la lecture par rapport à la rédaction du résumé en recourant à la chronologie par changement d'aspect et je dirai : après avoir achevé sa lecture il en rédige un résumé. Si je rejette l'action de rédiger dans un passé lié au présent, je l'énoncerai au passé composé : il en a rédigé un résumé. L'action de rédiger se situe maintenant au deuxième plan. L'action d'achever la lecture, plus ancienne, devrait donc glisser au troisième plan. Mais, précisément, ce décalage n'a pas lieu et toute la phrase s'énonce de la façon suivante : après avoir achevé sa lecture il en a rédigé un résumé.

Cornu (1953 : p. 145)

En résumé, les seules formes surcomposées que nous considérons correctes en français standard contemporain sont le passé surcomposé (*a eu fait*), le plus-que-parfait surcomposé (*avait eu fait*), le conditionnel surcomposé à valeur modale (*aurait eu fait*) et le futur antérieur surcomposé à valeur modale (*aura eu fait*). Les deux premières sont beaucoup plus fréquentes que les deux dernières, en particu-

lier le passé surcomposé, qui est de loin la forme la plus fréquente de toutes. Nous allons maintenant voir d'où proviennent ces formes et quelle en est la structure.

4.2 L'origine des formes surcomposées

La première publication d'importance dédiée aux formes surcomposées, à notre connaissance, a été celle de Foulet (1925), qui s'est attaché à retracer l'origine et l'évolution du «passé surcomposé» et du «plus-que-parfait surcomposé» (les formes surcomposées construites sur la base, respectivement, du passé composé [*il a eu mangé*] et du plus-que-parfait [*il avait eu mangé*]). Il estime que le passé surcomposé serait apparu dès le XIII^e siècle, bien qu'il ne soit clairement attesté dans les textes qu'à partir du milieu du XV^e siècle, alors que le plus-que-parfait surcomposé est déjà clairement attesté au XIII^e siècle. C'est à peu près à cette époque, que le passé simple aurait commencé à céder peu à peu le terrain au passé composé, à l'origine rien d'autre qu'un parfait³⁵. Ce serait donc en réponse au glissement sémantique des formes composées (du parfait vers le prétérit) que se serait répandu l'usage des surcomposés, afin de bien marquer l'aspect révolu de faits passés. C'est du moins la thèse la plus généralement acceptée, bien qu'elle ne fasse pas l'unanimité. Notamment, Grevisse (1993) est plutôt d'avis que le recul du passé simple est postérieur à l'apparition des surcomposés, aussi préfère-t-il y voir un simple mécanisme servant à marquer l'accompli par rapport aux formes composées. Il n'explique cependant pas pourquoi on aurait créé le passé surcomposé pour remplacer le passé antérieur, alors que les deux avaient exactement le même sens. Il semble que ce soit en fait le plus-que-parfait surcomposé qui soit apparu en premier pour marquer l'accompli par rapport au plus-que-parfait (ce qu'aucune forme verbale ne pouvait exprimer auparavant), et que le passé surcomposé ait été créé par analogie avec cette nouvelle forme. Cette hypothèse coïncide d'ailleurs avec les observations de Foulet (1925), qui date l'apparition du plus-que-parfait surcomposé dans les textes environ deux siècles avant celle du passé surcomposé.

35. Au XV^e siècle, le passé composé à valeur de prétérit était déjà courant dans les textes, et on peut supposer qu'il lui a fallu un certain temps pour être admis à l'écrit.

Le français est la seule des grandes langues romanes à connaître la surcomposition. Cependant, Cornu (1953) a relevé le phénomène dans plusieurs «patois» gallo-romans, gallo-italiens et rhéto-romans, des langues et dialectes qui auraient en commun avec le français d'avoir plus ou moins abandonné le passé simple, selon Cornu :

[...] les formes en question sont un phénomène caractéristique des idiomes qui ont abandonné peu à peu l'emploi du passé simple (de l'imparfait dans les langues germaniques) et du passé antérieur (du plus-que-parfait dans les langues germaniques) ou qui, tout au moins, ont fait du passé composé un rival équipollent du passé simple.

Cornu (1953 : p. 249)

Il juge vraisemblable que les formes surcomposées, dans leur usage «standard», se soient développées dans ces langues et dialectes sans grande influence mutuelle. On peut supposer que c'était pour résoudre le même problème, soit l'impossibilité de marquer l'aspect révolu de faits passés par des formes simples ou composées (et non surcomposées). Il croit cependant que les emplois dialectaux des surcomposés, puisqu'ils se sont retrouvés très tôt dans les textes écrits, ont pu passer d'une langue ou d'un dialecte à l'autre.

4.3 Surcomposé général vs surcomposé régional

La distinction entre ce qu'on appelle le **surcomposé général**, c'est-à-dire le surcomposé tel qu'on l'emploie en français standard, et le **surcomposé régional**, c'est-à-dire les usages du surcomposé dans certains dialectes du Sud de la France, est très importante. En effet, les usages dialectaux du surcomposé diffèrent passablement des emplois standard, non seulement dans leur sens, mais également parfois dans leur forme. Carruthers (1994) identifie cinq sens principaux que peut porter le passé surcomposé dans les dialectes du Sud mais qui sont exclus en français standard :

- 'le fait est définitivement complété et il est peu probable qu'il se reproduise';
- 'le fait a eu lieu dans un passé éloigné';
- 'le fait a eu lieu à une époque indéfinie ou indéterminée';

- ‘le fait est d’une certaine façon exceptionnel’;
- ‘le locuteur est directement impliqué dans le fait’.

Pour ce qui est de la forme, elle diffère dans les dialectes en question pour les verbes se conjuguant avec l’auxiliaire ÊTRE: alors qu’on a en français standard *il a été tombé*, par exemple, on trouve dans ces dialectes *il est eu tombé*. D’après Jolivet (1984), cela s’expliquerait par le fait que les formes surcomposées sont structurellement différentes dans ces dialectes. Rappelons d’abord la théorie de Tesnière (1939), selon laquelle les formes surcomposées sont construites par la composition de l’auxiliaire des formes composées, ce qui a pour conséquence que la coupure entre les deux auxiliaires est moins forte qu’entre le deuxième auxiliaire et le verbe, ce que Tesnière représentait ainsi: *a / eu // fait*. Jolivet note qu’il en découle une plus grande séparabilité du verbe et du second auxiliaire, ce qui se vérifie dans l’usage standard, où c’est entre *a eu* et le verbe qu’ont tendance à s’insérer les adverbes [*quand il a eu assez mangé*]. Or, il faut remarquer que cela ne se vérifie pas dans l’usage dialectal, où c’est plutôt entre les deux auxiliaires que s’insèrent les adverbes [*tu as bien eu joué pourtant*] (Jolivet, 1984: p. 172). Il en conclut que dans les dialectes en question, la coupure est *a // eu / fait*, c’est-à-dire que *eu fait* forment un bloc. Ce serait la raison pour laquelle on a *il est eu tombé* dans ces dialectes au lieu de *il a été tombé*. À notre avis, la plus grande cohésion qu’on observe en français standard entre *a* et *eu*, comparativement à *eu* et *fait*, est dû au fait que le premier relève de la flexion, mais pas le second, comme nous le verrons dans les sections suivantes (§ 5, p. 164 et § 6, p. 177).

Nous ne considérons dans notre travail que les emplois relevant du français standard et ignorons totalement ceux propres à certaines variétés du français³⁶, et à plus forte raison ceux qu’on peut trouver dans d’autres langues. Le lecteur intéressé par le surcomposé régional pourra consulter notamment les travaux de Foulet (1925), Cornu (1953), ou encore Carruthers (1993, 1994, 1999).

36. Carruthers (1998) montre que la distinction entre surcomposé général et surcomposé régional n’est pas toujours très nette. Elle reste cependant importante, notamment au niveau de leur comportement syntaxique.

4.4 La rareté des formes surcomposées

Ce qui complique l'étude des formes surcomposées est leur rareté. Une étude de Carruthers (1996, 1999) a permis notamment de mesurer cette rareté. Sur un corpus de 100 heures d'entrevue dirigée, on ne trouve que 14 occurrences de formes surcomposées relevant du français standard (contre 60 occurrences de formes surcomposées dialectales), alors que l'entrevue était justement conçue de manière à favoriser la production de ces formes. Au cours de la même étude, seulement 30% des adultes interrogés ont utilisé des formes surcomposées dans un «texte à trou» écrit. Dans un test similaire à l'oral, ce nombre tombait à seulement 12%. De plus, l'auteure signale que le passé surcomposé, c'est-à-dire la forme construite à partir d'un auxiliaire AVOIR au présent [*j'ai eu fait*] est beaucoup plus courant que les autres formes surcomposées. Ces données suggèrent que les formes surcomposées ne sont pas facilement accessibles aux locuteurs du français. Pourtant, dans le cadre d'un test d'acceptabilité réalisé lors de cette même étude, la majorité d'entre eux (71%) considéraient ces formes correctes.

Le fait que des formes grammaticales aussi rares soient si largement acceptées par les locuteurs suggère qu'il s'agit de formes compositionnelles, construites à partir de signes plus simples et plus fréquents, dont seule la combinaison serait rare. La seule chose qui soit à la fois propre seulement aux formes surcomposées et commune à toutes celles-ci est la présence de deux auxiliaires (chacun imposant le participe passé à son dépendant). Il faut maintenant déterminer s'il s'agit de deux occurrences d'un même signe qui se combine à lui-même ou si on a affaire à deux signes distincts.

4.5 Auxiliaire vague ou polysémique?

Ce qui caractérise les formes surcomposées est leur usage double de l'auxiliaire AVOIR³⁷. À notre connaissance, personne n'a encore cherché à déterminer de façon rigoureuse s'il s'agissait de deux occurrences d'un même auxiliaire se com-

37. Afin d'alléger le texte, nous faisons abstraction dans cette partie de l'alternance AVOIR ~ ÊTRE, puisqu'elle n'est pas pertinente pour notre propos.

binant à lui-même, ou de deux auxiliaires se combinant entre eux. Dans les ouvrages que nous avons consultés, on considère (souvent implicitement) qu'il s'agit d'un seul auxiliaire qui se combine avec lui-même, sans que la position ne soit justifiée. Pourtant, certains faits suggèrent le contraire.

Premièrement, remarquons qu'il existe des formes à deux auxiliaires AVOIR, mais jamais à trois ou plus³⁸:

(71) *j'ai fini*

(72) *j'ai eu fini*

(73) * *j'ai eu eu fini*

Si on avait affaire à un auxiliaire se combinant avec lui-même, on pourrait s'attendre à pouvoir en aligner un nombre en théorie illimité d'occurrences, comme on peut le faire par exemple avec le préfixe RE- [*Je vous le rereredis*]. Nous voyons mal comment expliquer de façon élégante cette restriction sur la combinatoire d'un éventuel auxiliaire AVOIR unique.

Deuxièmement, il est généralement admis, au moins depuis le célèbre article de Benveniste (1959), que le passé composé peut exprimer soit l'antériorité (il commute alors aisément avec le passé simple et est compatible avec un complément «*en + durée*»), soit l'aspect accompli³⁹ (auquel cas la substitution n'est pas possible, et on peut utiliser un complément «*depuis + durée*»):

(74) a. *J'ai terminé le tableau en deux heures.*

b. *Je terminai le tableau en deux heures.*

c. *J'ai terminé le tableau depuis deux heures.*

d. * *Je terminai le tableau depuis deux heures.*

38. Cornu (1953) fait état de formes «hypercomposées» à trois auxiliaires AVOIR dans certains patois de la Suisse romande, par exemple *Mon père-gran a zau zu improntaon capitô* 'mon grand-père a emprunté (du temps où il vivait encore) un capital' (*ib.* p. 225). Ces formes s'expliquent par le fait qu'elles sont construites sur la base de formes surcomposées non compositionnelles, c'est-à-dire que la construction à deux auxiliaires est figée.

39. Sur cette double valeur du passé composé, voir également Imbs (1960), Curat (1991) ou encore Leeman-Bouix (1994).

La phrase (c) peut être paraphrasée en utilisant l'expression *se trouver avoir V-é*⁴⁰, qui fait ressortir le caractère actuel (et non passé) des formes marquant l'accompli :

- (75) a. *Je me trouve avoir terminé le tableau depuis deux heures.*
 b. * *Je me trouve avoir terminé le tableau en deux heures.*

Les formes de l'accompli s'accommodent également très bien de l'adverbe MAINTENANT, ce qui n'est pas le cas des formes qui marquent l'antériorité :

- (76) a. *Maintenant j'ai terminé le tableau.*
 b. * *Maintenant je terminai le tableau.*

Il y a donc clairement deux sens associés à l'auxiliaire au passé composé. Nous reviendrons plus loin (§ 5, p. 164 et § 6, p. 177) sur ces sens, contentons-nous pour l'instant de ces descriptions sommaires. Ce qui nous intéresse ici, nous le rappelons, est de déterminer si ces deux sens correspondent à deux acceptions d'un vocable polysémique, ou si nous avons affaire à un seul signe au sens vague. C'est ici que nous sont utiles les formes surcomposées, puisqu'elles portent justement ces deux sens. On peut le vérifier en paraphrasant un des auxiliaires AVOIR de la phrase (77) ci-dessous par un passé simple et l'autre par l'expression *se trouver avoir V-é*, ou encore en utilisant le passé antérieur :

- (77) *Dès que j'ai eu terminé ma soupe, je suis sorti.*
 (78) *Dès que je me trouvai avoir terminé ma soupe, je sortis.*
 (79) *Dès que j'eus terminé ma soupe, je sortis.*

40. Cette construction est un peu lourde et peut-être vieillie, mais elle est attestée. Par exemple :

Osera-t-elle se remonter aux yeux d'un mari qu'elle adore, qu'elle se trouve avoir trompé sans qu'il y ait de sa faute et qu'elle verra mourir de chagrin d'avoir pu songer à la prendre pour femme? (Marquis de Sade, *La marquise de Telème ou les effets du libertinage*, 1788, texte en ligne). Ou encore : *L'abhumanisme serait la disponibilité absolue, le refus de sanctionner le choix qu'on se trouve avoir fait, ou avoir subi, dans l'insondable passé, d'être un homme [...]* (Audiberti, *L'ouvre-boîte*, 1952, texte en ligne). On la retrouve également au passé simple, où elle insiste sur le début de la phase dénotée par l'accompli : *C'est ainsi que Tchartkov se trouva avoir acheté un vieux portrait, en pensant à part soi : « Pourquoi l'ai-je acheté? Qu'est-ce que je vais en faire? »* (Gogol, *Le portrait*, 1835, texte en ligne, traducteur inconnu).

Si on admet que les surcomposés se construisent avec deux lexies AVOIR, chacune apportant un sens distinct, on voit très bien comment se compose le sens de ces constructions.

D'ailleurs, la phrase (78) ci-dessus suggère que les deux auxiliaires des formes surcomposées n'opèrent pas sémantiquement en bloc, mais que chacun est associé à une composante précise du sens de ces formes. On peut identifier exactement quel auxiliaire apporte quelle partie de ce sens, grâce à l'insertion de *se trouver*, qui permet en quelque sorte de «marquer» l'auxiliaire de l'accompli afin de pouvoir distinguer les deux AVOIR. Ainsi, on voit en (78) que *j'ai eu terminé* dénote un état ('accompli') qui a eu lieu dans le passé ('antérieur'), puisque *se trouver* porte le suffixe du passé simple. On peut alors en conclure que le passé surcomposé indique un «aoriste du parfait», et non un «parfait de l'aoriste» comme le soutenait Benveniste (1959).

4.6 Les formes surcomposées et la voix

Dans la liste des formes surcomposées que nous avons donnée ci-dessus (§ 4.1, p. 145) nous avons fait abstraction de la voix. Nous allons maintenant voir ce qu'il en est de l'usage des formes surcomposées aux autres voix. Rappelons que nous considérons dans notre modèle, outre la voix active, trois voix passives et deux voix réfléchies (nous reprenons ici les exemples donnés plus haut, § 2, p. 132):

- **Passif₁**: *Alain a été rasé par Jean.*
- **Passif₂**: *Une barbe de deux jours se rase sans difficulté.*
- **Passif₃**: *Il a été procédé par le gouvernement au licenciement des fonctionnaires inutiles.*
- **Réfléchi₁**: *Alain se rase.*
- **Réfléchi₂**: *Alain s'achète une maison. Alain se rase la barbe.*

Les **passif₁** et **passif₃** correspondent au passif de la terminologie traditionnelle, alors que les **passif₂**, **réfléchi₁** et **réfléchi₂** correspondent à ce qu'on appelle habituellement la voix pronominale.

Les formes surcomposées passives (au sens traditionnel du terme) étant rares, on a tendance à les considérer avec suspicion :

On ne constate, dans la liste de nos exemples, aucun cas de forme surcomposée passive. C'est que le passif est déjà, dans les formes composées du passé, un surcomposé (cf. cela a été entendu); le surcomposé en ferait un supersurcomposé, à quoi la langue semble répugner⁴¹.

Imbs (1960: p. 134)

Pourtant, elles ne sont pas forcément agrammaticales :

- (80) *Dès qu'Alain a eu été rasé par Jean, Hélène l'a embrassé. (passif₁)*
 (81) *Dès qu'il a eu été procédé par le gouvernement au licenciement des fonctionnaires inutiles, les problèmes ont commencé. (passif₃)*

Il est vrai cependant que les phrases qui en résultent peuvent sembler un peu lourdes, mais ce n'est pas à notre avis à cause du nombre d'auxiliaires qu'elles comptent. Il semble raisonnable effectivement de croire que plus on ajoute d'auxiliaires, moins la phrase est élégante, mais le français peut certainement tolérer la présence de trois auxiliaires. Par exemple :

- (82) *Les condamnés vont déjà avoir été exécutés à notre arrivée.*
 (83) *Les condamnés vont être sur le point de se faire fusiller à notre arrivée.*

Le problème vient peut-être simplement du fait que le passif n'est pas particulièrement utilisé en français et qu'il peut parfois sembler peu naturel, même au présent de l'indicatif. Sa combinaison avec le passé surcomposé, un temps encore plus rare et assez fortement stigmatisé, en fait certainement reculer plusieurs. De fait, on semble lui préférer, à l'écrit, le passé antérieur passif, même dans un texte où on n'utilise pas le passé simple, comme on peut l'observer dans l'exemple suivant (remarquez le passé composé dans la deuxième partie de la phrase) :

- (84) *Après qu'il eut été déclaré inconstitutionnel en août dernier, le programme d'écoutes électroniques sans mandat judiciaire du président George W. Bush a reçu l'approbation de la Chambre des représentants, jeudi.*
 www.radio-canada.ca, 29 septembre 2006

41. Imbs adopte la définition de Tesnière (1939) du terme «surcomposé», qui désigne pour eux toute forme comportant deux auxiliaires, sans égard à la nature de ces auxiliaires. Nous utilisons plutôt la définition la plus courante du terme, selon laquelle seules les formes à deux auxiliaires AVOIR sont dites surcomposées (cf. § 4.1, p. 145).

Les trois autres voix de notre modèle, qui correspondent à la voix pronominale dans la tradition, sont aussi sujettes à la surcomposition⁴²:

- (85) *Dès que le gâteau s'est eu tout mangé, Rita en a fait un deuxième.* (**passif₂**)
 (86) *Dès qu'Alain s'est eu rasé, Hélène l'a embrassé.* (**réfléchi₁**)
 (87) *Dès qu'Alain s'est eu rasé la barbe, Hélène l'a embrassé.* (**réfléchi₂**)

Les verbes essentiellement pronominaux ne se comportent pas différemment, mais il faut noter que ces constructions présentent une particularité importante, comme le remarquait entre autres Grevisse:

*Les verbes pronominaux ne présentent pas les formes attendues [...]: Quand je me suis assis (passé composé) → *Quand je m'ai été assis, mais → Quand je me suis eu assis (passé surcomposé). Il y a comme une permutation des auxiliaires, à cause du fait que le pronom réfléchi ne peut pas être suivi de l'auxiliaire avoir.*

Grevisse (1993: § 788)

Cela contraste avec les verbes qui se conjuguent avec ÊTRE mais qui ne sont pas pronominaux:

- (88) *Quand il a été parti, les enfants sont sortis de leur cachette*⁴³.

Encore une fois, ces faits nous portent à croire qu'en français standard, le second auxiliaire des formes surcomposées n'est pas d'ordre flexionnel alors que le premier l'est. Tout se passe comme si le second auxiliaire devenait la nouvelle base de la construction pronominale, c'est-à-dire que le passé composé dans *J'ai déjà mangé* n'est pas une forme fléchie de MANGER au sens strict. C'est la position que nous défendons dans les sections suivantes (§ 5, p. 164 et § 6, p. 177).

42. Une étude de Carruthers (1999) suggère cependant que les verbes pronominaux surcomposés sont d'un emploi très rare (son corpus n'en a révélé aucune occurrence, contre 14 occurrences de verbes non-pronominaux surcomposés). Une requête pour «s'est eu» sur Google nous a retourné un nombre anormalement élevé de sites et de livres qui traitent des surcomposés, ce qui est assez symptomatique. Nous avons néanmoins trouvé cet exemple authentique: *Quand il s'est eu calmé je lui ai bien expliqué pourquoi je peux pas le laisser se comporter comme ça* (blog *Matahari's new's* [sic]).

43. Dans les dialectes du Sud de la France, au contraire, les verbes qui se conjuguent avec ÊTRE présentent l'ordre inverse [(rég.) *Quand il est eu parti...*].

4.7 Synthèse

Les formes surcomposées nous révèlent qu'il existe en français deux auxiliaires **AVOIR** qui ne sont pas toujours interchangeables: un qui exprime l'antériorité (nous l'appellerons dorénavant **AVOIR_{ant}**) et un autre qui exprime l'aspect accompli (que nous appellerons **AVOIR_{acc}**). On notera que cette ambiguïté de **AVOIR** se trouve tant dans le plus-que-parfait que dans le passé composé, comme le révèlent les tests «*en + durée*» et «*depuis + durée*»:

- (89) a. *J'ai terminé le tableau en deux heures.*
 b. *J'ai terminé le tableau depuis deux heures.*
- (90) a. *J'avais terminé le tableau en deux heures.*
 b. *J'avais terminé le tableau depuis deux heures.*

On a pu également déterminer que le temps avait portée sur l'aspect, et non l'inverse, c'est-à-dire que les formes surcomposées sont formées de la séquence «**AVOIR_{ant} + AVOIR_{acc} + V**».

Nous tenterons dans les sections qui suivent de cerner la différence sémantique entre ces deux auxiliaires et de déterminer de façon plus précise la nature de leur sens, en commençant par l'accompli.

5 Les phases aspectuelles

Nous venons de voir qu'il existe deux auxiliaires **AVOIR**: un qui exprime l'accompli (**AVOIR_{acc}**) et un autre qui exprime l'antériorité (**AVOIR_{ant}**). Nous les considérons comme des lexies distinctes qui peuvent être étudiés séparément, bien qu'elles soient étroitement liées. Nous allons d'abord nous pencher sur **AVOIR_{acc}** ainsi que sur les lexies qui lui sont apparentées, et nous reviendrons sur **AVOIR_{ant}** et ses semblables à la section suivante (§ 6, p. 177). Nous nous intéresserons dans la présente section principalement aux formes construites avec les auxiliaires **AVOIR_{acc}** ou **ALLER**, ou encore avec la locution prépositive **EN TRAIN**⁴⁴.

44. Il est possible que *en train* soit un seul mot-forme. La question ne nous intéresse pas ici.

Ces signes ont notamment en commun d'être souvent regroupés dans la littérature sous le terme d'«aspect». Le terme semble approprié, puisque les signes en question font tous référence d'une façon ou d'une autre à la structure temporelle interne des faits en ce qu'ils insistent sur le début, la continuation ou la fin de ces faits. Cependant, on regroupe également sous le terme d'«aspect» d'autres phénomènes linguistiques qu'il ne faut pas confondre avec ceux à l'étude ici, notamment la perfectivité (nous reviendrons plus loin sur ce type d'aspect, § 6.1.2, p. 181). Nous préférons donc utiliser le terme «aspect» dans un sens générique qui englobe un éventail de phénomènes linguistiques. Pour les signes qui nous intéressent ici, nous utilisons le terme de «phase aspectuelle», dont nous allons tout de suite préciser le sens.

5.1 Les phases naturelles et leur manifestation en français

Les faits dénotés par les verbes ont une certaine durée (éventuellement presque nulle dans le cas des faits ponctuels) et par conséquent ils divisent le temps réel en trois phases naturelles: l'avant, le pendant et l'après.

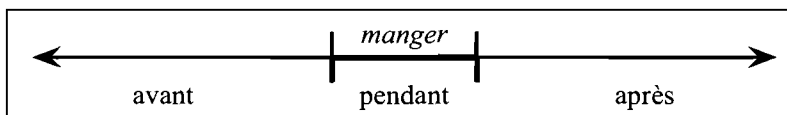


Figure 20 — Le découpage du temps réel en trois phases par un fait

D'après Wagner & Pinchon (1962), cette réalité se reflète dans le système grammatical du français, et ils décrivent les auxiliaires, semi-auxiliaires et locutions prépositives en fonction de ces phases et des bornes qui les délimitent. Ce que nous appelons «**phase aspectuelle**»⁴⁵ est une catégorie flexionnelle hypothétique dont les grammèmes serviraient à refléter ces phases naturelles. Nous nous proposons d'étudier sous cet angle les signes qui nous intéressent ici, puis nous discuterons le statut de cette catégorie flexionnelle hypothétique.

45. Afin d'alléger le texte, nous utilisons aussi parfois le terme «phase» tout court.

5.2 L'accompli

Pour Wagner & Pinchon (1962), l'auxiliaire **AVOIR_{acc}** dénote la phase qui suit un fait. Cela rejoint l'idée largement consensuelle selon laquelle le passé composé situe un fait dans le passé tout en ayant un lien avec le présent, voire tout en étant carrément un présent. En fait, l'accompli ne situerait pas vraiment un fait dans le passé, mais il dénoterait plutôt un état tout à fait contemporain. Cet état pourrait se définir comme suit: 'X a Y-é' = 'X se trouve dans l'état qui suit le fait Y'⁴⁶. Cette définition n'est toutefois pas adéquate, puisqu'il est possible de dire *Il a plu*, situation qui n'implique pas d'actant. Nous allons donc plutôt définir la phase accomplie ainsi: 'avoir X-é' = 'la période qui suit le fait X a cours'. Ce n'est que par inférence qu'on déduit que le fait dénoté par le verbe s'est déroulé dans le passé, mais cette information ne fait pas partie du sens de l'accompli, comme l'a si bien formulé Leeman-Bouix (1994) à propos de la phrase *Il faut que j'aie fait les courses à 3 heures*:

L'interprétation «passé» vient de ce que, si les courses sont présentées comme faites à 3 heures, on en déduit qu'elles ont été faites avant 3 heures, donc, si l'on veut, dans le passé par rapport à 3 heures; mais cette conséquence interprétative ne nous autorise pas à définir aie fait comme exprimant le passé.

Leeman-Bouix (1994: p. 56)

On retrouve la même idée chez Mel'čuk (1994), qui explique que «le pont entre le temps et l'aspect est assuré par 'accompli', qui, psychologiquement, implique l'idée de 'passé'» (p. 51). Nous sommes d'accord avec eux et croyons que l'accompli ne présente pas un fait comme étant avant maintenant, mais plutôt maintenant comme étant après un fait. Cette différence de perspective est à notre avis ce qui distingue l'accompli du passé. Nous y reviendrons dans notre discussion du temps passé (§ 6.4.2, p. 205).

Selon la position relative de la phase accomplie par rapport au moment d'énonciation (qui peut se trouver dans cette phase ou non), l'auxiliaire **AVOIR_{acc}** peut

46. C'est aussi ce que semblent vouloir exprimer Wagner & Pinchon (1962: p. 348) quand ils disent que *J'ai vu* équivaut à *Je suis dans la situation de quelqu'un qui a vu*, mais la circularité de leur définition la rend inexploitable.

être fléchi de diverses façons. Par exemple, si au moment où j'utilise la construction «AVOIR_{acc} + mangé» je me trouve dans la phase qui suit le fait de manger, alors la phase dénotée par «AVOIR_{acc} + mangé» est simultanée au moment d'énonciation (que nous symbolisons par T_0), ce qu'on peut représenter visuellement comme ceci :

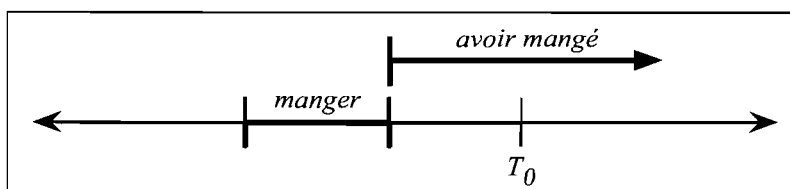


Figure 21 — La phase accomplie est simultanée au moment d'énonciation

Dans cette situation, l'auxiliaire AVOIR_{acc} doit absolument être au présent :

(91) *Il a mangé.*

Autrement, si la phase dénotée par l'accompli ne commence qu'après le moment d'énonciation, il faut alors utiliser une des formes suivantes :

(92) *Il aura mangé.*

(93) *Il va avoir mangé.*

Pour ces phrases, la position relative du fait de manger par rapport au moment d'énonciation n'est pas précisée. Seule la position relative de la phase qui suit le fait de manger par rapport au moment d'énonciation est pertinente. Ces phrases peuvent donc correspondre indifféremment aux deux situations suivantes :

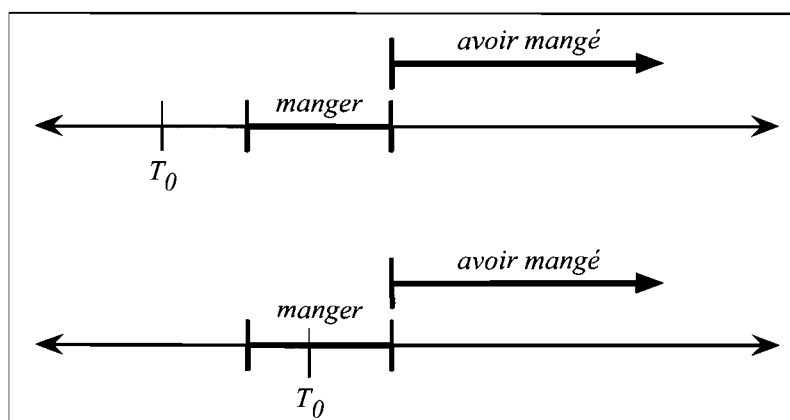


Figure 22 — La phase accomplie est postérieure au moment d'énonciation

Comme on peut s'y attendre, si la phase se trouve avant le moment d'énonciation, l'auxiliaire devra avoir une forme qui dénote un passé, en se combinant soit à la construction «AVOIR_{ant} + V_{pp}» (qui marque l'antériorité), soit au morphème du passé simple ou de l'imparfait :

(94) *il a eu mangé*

(95) *il eût mangé*

(96) *il avait mangé*

Il y a toutefois un problème ici. Étant donné que nous avons défini la phase accomplie comme étant celle qui suit un fait, elle n'a pas de borne finale (ce que nous avons représenté visuellement par des vecteurs). Mais comment cette phase qui n'a pas de fin peut-elle se trouver avant le moment d'énonciation? Nous verrons plus loin (§ 6, p. 177) que le plus-que-parfait ne pose pas de problème puisqu'il indique en fait que la phase dénotée par AVOIR_{acc} est simultanée à un point de référence passé (elle n'est donc pas directement présentée comme antérieure au moment d'énonciation). Ce n'est toutefois pas le cas des deux autres formes, qui situent cette directement par rapport au moment d'énonciation. Alors, comment expliquer la contradiction apparente pour ces formes?

En fait, le problème est limité à quelques constructions. En effet, ces formes ne s'utilisent guère que dans des contextes bien particuliers, notamment en combinaison avec QUAND, 「DÈS QUE」, 「SITÔT QUE」, etc. :

(97) *Dès qu'il a eu mangé sa soupe, il est parti.*

(98)* *Il a eu mangé sa soupe.*

Il s'agit des contextes d'utilisation des formes surcomposées (§ 4, p. 144). Or, ces contextes semblent avoir une certaine influence sur la façon dont les sens se combinent. Considérons par exemple la phrase suivante :

(99) *Dès que Jean fut beau, il cessa de fumer.*

Le sémantème 'beau' dénote certainement un état qui n'est pas linguistiquement borné, comme en témoigne l'agrammaticalité des phrases suivantes :

(100)**Il vient d'être beau.*

(101)**Il est sur le point d'être beau.*

Pourtant, en (99), *être beau* acquière un sens inchoatif et peut être situé dans le passé. On comprend alors que la personne en question a fait ou subi quelque chose qui l'a soudainement rendue belle. Ce sens provient de la construction dans laquelle apparaît cette expression. Nous croyons que c'est le même phénomène qui se produit pour le passé surcomposé et le passé antérieur. L'auxiliaire AVOIR_{acc} dénote bien un fait qui n'a pas de borne finale (et qui ne peut donc pas en principe être situé directement dans le passé par rapport au moment d'énonciation), mais QUAND, 「DÈS QUE」, etc. dénotent le moment où cette phase commence, et c'est alors ce moment qui est situé dans le passé, soit par l'addition de AVOIR_{ant} (pour le passé surcomposé), soit par l'ajout du passé simple (pour le passé antérieur).

5.3 Le progressif

Le meilleur candidat pour dénoter la phase de continuation d'un procès (que nous appellerons la «phase progressive») est sans conteste la construction 「EN TRAIN」⁴⁷. Nous avons déjà démontré (Ch. II, § 3, p. 28) qu'il s'agit d'une locution prépositive et non d'un auxiliaire ou semi-auxiliaire. En effet, la copule ÊTRE qui l'accompagne le plus souvent n'est pas une composante de cette construction. Cette locution ne peut donc pas faire partie du système verbal du français au sens strict puisqu'elle ne relève pas de la classe des verbes, mais bien de celle des prépositions. Elle est aux verbes ce que 「EN PLEIN」 [*Ils sont arrivés en plein repas*] est aux noms. Nous allons néanmoins tenir compte de ce signe et le modéliser au Chapitre VIII puisqu'il est pertinent pour le système verbal du français, étant donné qu'il dénote une des trois phases naturelles délimitées par les faits. Il sera également intéressant de voir comment une locution prépositive peut se fondre dans le système verbal (cf. Ch. VIII, § 3.3.2, p. 328). On peut définir provisoirement cette locution comme suit: 'en train de X-er' = 'la période simultanée au fait X a cours'. Comme pour l'auxiliaire de l'accompli, la position relative de la phase progressive

47. L'expression «ALLER + V-ant», comme dans *Le mal va croissant* (cf. Riegel, Pellat & Rioul 1994: p. 296) est trop archaïque.

par rapport au moment d'énonciation détermine la flexion que porte la copule ÊTRE, si 「EN TRAIN」 se trouve dans une position syntaxique qui exige cette copule.

5.4 Le prospectif

Enfin, la dernière phase, celle qui précède le fait, est en quelque sorte le miroir de l'accompli. Nous l'appellerons «phase prospective». De la même façon que l'accompli a un sens qui ressemble beaucoup à un passé sans en être un, cette phase doit avoir un sens qui ressemble à un futur mais qui présente le moment d'énonciation comme précédant le fait, et non le fait comme postérieur au moment d'énonciation. Il y a deux candidats comme marqueurs potentiels de cette phase: le futur simple et la construction «ALLER + V_{inf}». Nous avons vu que pour les deux autres phases, lorsque le moment d'énonciation est simultanément à la phase on a formellement un auxiliaire fléchi au présent. C'est bien ce à quoi on peut s'attendre, puisque si le locuteur se trouve dans la phase dont il parle, cette phase est un fait actuel, donc le présent s'impose. L'auxiliaire ALLER semble donc être le meilleur candidat pour le poste à combler. Cette hypothèse est encore renforcée par le fait que, comme l'accompli, l'auxiliaire ALLER s'accommode très bien d'adverbes tels que MAINTENANT et ses semblables. On comparera les phrases (76) a) et (b) à celles-ci (adaptées de Leeman-Bouix 1994: p. 163):

(102) *Je vais l'appeler tout de suite.*

(103)**Je l'appellerai tout de suite*⁴⁸.

La tradition (et avec elle un bon nombre d'auteurs contemporains)⁴⁹ considère la construction «ALLER + V_{inf}» comme marquant le «futur proche», miroir de VENIR, au même titre que 「SUR LE POINT」:

Dans des phrases comme Le vase va se casser ou Le vase vient de se casser, le verbe aller traduit un futur proche et venir de un passé proche: ils marquent respectivement l'immédiateté future ou passée.

Leeman-Bouix (1994: p. 119)

48. Dans *Je l'appellerai tout de suite en rentrant*, il s'agit d'une autre acception de 「TOUT DE SUITE」.

49. Notamment Wagner & Pinchon (1962), Riegel, Pellat & Rioul (1994), Leeman-Bouix, (1994) ou encore Gosselin (1996), pour n'en nommer que quelques-uns.

Or, la phrase *Dans cinq milliards d'années, le soleil va commencer à grossir* est tout à fait correcte, ce qui démontre que le sens de l'auxiliaire ALLER n'inclut aucune idée de proximité temporelle. Ce genre de phrase n'est pas possible, par exemple, avec 'SUR LE POINT', qui dénote la période qui précède immédiatement un fait [**Dans cinq milliards d'années, le soleil est sur le point de commencer à gonfler*] (cf. § 5.5, p. 172).

Par ailleurs, l'auxiliaire ALLER n'a pas la même combinatoire que les constructions qui marquent vraiment le futur proche. Il semble en effet s'accommoder d'un verbe qui dénote un état⁵⁰, alors que ces constructions sont incompatibles avec de tels verbes :

(104) *Elle va être enceinte.*

(105)**Elle est sur le point d'être enceinte.*

(106)*(québ.) *Elle s'en allait être enceinte.*

Cette propriété est d'ailleurs partagée par l'accompli :

(107) *Il a vécu heureux.*

(108)**Il vient de vivre heureux.*

Nous considérons donc que ALLER n'est pas un marqueur du futur proche, symétrique à VENIR, mais plutôt un marqueur de la phase prospective, symétrique à AVOIR_{acc}. Nous pouvons en décrire le sens ainsi : 'aller X-er' = 'la période qui précède le fait X a cours'.

Enfin, notons que ALLER ne se combine qu'à l'indicatif, et encore, seulement à l'imparfait ou au présent, alors que tous les marqueurs dénotant des phases ou des bornes temporelles (ou leur région immédiate) peuvent s'utiliser à n'importe quel mode, temps et finitude [*être sur le point de finir, venant d'arriver, aura bu, commencerait à pleuvoir, finira de compter, etc.*].

50. Nous présupposons que la copule hérite des propriétés sémantiques de son complément.

5.5 L'expression des bornes temporelles des faits

Les bornes temporelles des faits ayant elles-mêmes une certaine saillance d'un point de vue cognitif, on peut aussi s'attendre à ce qu'elles fassent, avec leur périphérie immédiate, l'objet d'une attention particulière qui se cristallise dans la langue. De fait, la littérature relève un certain nombre de signes linguistiques verbaux ou «pseudo-verbaux» (des locutions prépositives utilisées avec ÊTRE) qui font directement référence à ces bornes. Voici une liste non exhaustive des signes dont nous avons trouvé la mention dans diverses grammaires du français: 「SUR LE POINT」, 「À LA VEILLE」, 「EN PASSE」, 「EN VOIE」, 「SE METTRE」, COMMENCER, ENTREPRENDRE, ARRÊTER, FINIR, TERMINER, CESSER, VENIR. On peut encore y ajouter les collocatifs qui n'ont de tels sens que dans le contexte de certains verbes, par exemple *partir* dans (québ.) *partir à rire* [= *Incep*(RIRE)]⁵¹. Toutes les locutions prépositionnelles sont d'un intérêt limité pour nous, puisqu'elles ne peuvent en aucun cas faire partie d'une catégorie flexionnelle verbale. Nous avons déjà mentionné les cas de 「EN TRAIN」, qui nous suffira à illustrer au Chapitre VIII la mécanique qui permet la combinaison de ces locutions prépositives avec les verbes. Quant aux verbes de notre liste, presque tous peuvent se combiner autant avec un nom qu'avec un verbe, ce qui en fait de mauvais candidats pour une éventuelle catégorie flexionnelle. En outre, ils ont presque tous des propriétés syntaxiques qui les distinguent des vrais auxiliaires, par exemple leur capacité à apparaître sans leur verbe dépendant:

(109) – *Est-ce qu'il a cessé de tourner?*
– *Non, il n'a pas encore cessé.*

(110) – *Est-ce que tu as mangé?*
– * *Oui, j'ai* < * *Oui, je l'ai* >.

Par ailleurs, ces verbes peuvent aussi être combinés à des affixes dérivationnels, ce qui n'est pas possible avec les auxiliaires:

51. Les verbes n'ont pas tendance à prendre des collocatifs verbaux, probablement parce qu'ils peuvent être le sommet syntaxique d'une proposition, ce qui rend ce type de collocation beaucoup plus rare que, par exemple, la combinaison d'un collocatif verbal et d'une base nominale [*prendre peur, attraper froid*, etc.].

(111) *Il se remet à boire.*

(112)* *Il re-a bu <* Il re-va boire>.*

Rien ne semble indiquer que ces verbes aient quelque chose de particulier qui les distingue de n'importe quel autre verbe. La seule exception est VENIR, qui est manifestement un auxiliaire :

(113) – *Est-ce qu'il a fini de manger?*
– * *Oui, il (le) vient.*

(114)* *Il revient de partir.*

Nous pourrions en définir provisoirement le sens ainsi: 'venir de X-er' = 'la phase qui suit immédiatement le fait X a cours'. D'après nos définitions, le sens de AVOIR_{acc} est entièrement inclus dans celui de VENIR, ce qui veut dire que toute situation pouvant être décrite par VENIR peut aussi l'être, de façon moins précise, par AVOIR_{acc}.

Son miroir est 'SUR LE POINT', qui contient le sens de ALLER, et que nous définissons provisoirement ainsi: 'sur le point de X-er' = 'la phase qui précède immédiatement le fait X a cours'.

5.6 Y a-t-il une catégorie flexionnelle de phase aspectuelle?

De toutes les lexies que nous avons examinées dans cette section, seulement trois sont des auxiliaires: AVOIR_{acc}, ALLER et VENIR. Ces auxiliaires présentent une parenté sémantique évidente et, avec les locutions 'EN TRAIN' et 'SUR LE POINT', ils semblent former un système qui offre une belle symétrie d'un point de vue sémantique, ce qui a sans doute poussé Wagner & Pinchon (1962) à les présenter ensemble dans leur grammaire. Pourtant, ces signes ne peuvent pas former une catégorie flexionnelle au sens où nous l'entendons (*cf.* Ch. IV, § 6, p. 62). D'une part, comme nous l'avons déjà mentionné, les locutions prépositives ne peuvent pas faire partie d'une catégorie flexionnelle verbale. D'autre part, même en faisant abstraction de ce problème, il reste que ces marqueurs ne sont pas tout à fait mutuellement exclusifs, bien que la plupart des combinaisons soient exclues. Le

tableau ci-dessous présente la combinatoire de ces marqueurs (nous reprenons ici en partie le Tableau II, p. 31):

	AVOIR _{acc}	「EN TRAIN」	ALLER	VENIR	「SUR LE POINT」
AVOIR _{acc}	*	* ⁱ	*	*	* ⁱⁱ
「EN TRAIN」	*	*	*	*	*
ALLER	✓	✓	*	✓	✓
VENIR	*	*	*	*	*
「SUR LE POINT」	? ⁱⁱⁱ	*	*	*	? ^{iv}

Tableau XVI — La combinatoire des marqueurs de phase

- i. *Quand il a été en train de manger* n'est pas agrammatical, mais il s'agit de AVOIR_{ant} (qu'on peut remplacer par le passé simple: *Quand il fut en train de manger*).
- ii. *Quand il a été sur le point de partir* est également construit avec AVOIR_{ant} (cf. *Quand il fut sur le point de partir*).
- iii. *Il est sur le point d'avoir fini* est plutôt lourd mais pas inusité, comme en témoigne cette phrase trouvée sur internet: *Une petite tradition du lieu: lorsque vous êtes sur le point d'avoir fini votre verre, le patron vous en remet un direct.*
- iv. ? *Il est sur le point d'être sur le point de partir* peut à la limite être accepté comme une plaisanterie.

ALLER se combine librement aux autres marqueurs. Il ne peut donc pas faire partie de la même catégorie flexionnelle que les autres auxiliaires. Mais qu'en est-il de VENIR et AVOIR_{acc}, qui sont des verbes et qui sont mutuellement exclusifs? Nous croyons que l'impossibilité de les combiner n'est pas une conséquence de leur appartenance à une même catégorie flexionnelle, mais qu'elle est due à des facteurs sémantiques. Si les grammèmes d'une même catégorie flexionnelle sont mutuellement exclusifs, c'est parce qu'ils doivent se combiner au même hôte, alors qu'ils portent des informations contradictoires. Par exemple, le singulier et le pluriel des noms sont mutuellement exclusifs parce qu'il est impossible de quantifier

un même nom à la fois au singulier et au pluriel. Le cas des marqueurs de phase est différent. Dans l'expression * *il vient d'avoir dormi*, le sens de VENIR ne porte pas sur celui de DORMIR, mais bien sur celui de AVOIR_{acc}. Si cette expression est agrammaticale, c'est simplement parce que ces deux auxiliaires sont sémantiquement incompatibles. En effet, d'après les définitions que nous avons données pour les marqueurs de phase, le sens de ces signes contient une référence indirecte aux bornes des faits (par les composantes sémantiques 'avant' et 'après'). C'est ce qui explique l'incompatibilité de ces signes avec des verbes dénotant des faits qui ne sont pas linguistiquement bornés⁵²:

(115)* *Il est sur le point d'être mort.*

(116)* *Il vient de vivre*⁵³.

Donc, le fait que les marqueurs de phase ne soient pas eux-mêmes bornés linguistiquement explique l'impossibilité de les combiner entre eux:

(117)* *Je viens d'être en train de danser.*

(118)* *Ils étaient sur le point d'avoir gagné.*

Cela a pour effet que ces signes sont mutuellement exclusifs, ce qui les rapproche des grammèmes, mais la mécanique derrière ce phénomène est bien différente de celle qui empêche la combinaison mutuelle de vrais grammèmes. Les quelques cas présentés ici nous suffisent. Pour d'autres exemples, voir notamment Chu (2008).

Par ailleurs, les locutions prépositionnelles, en plus de ne pas être des verbes, peuvent être remplacées par un certain nombre de constructions synonymes dans un même dialecte ou d'un dialecte à l'autre [SUR LE POINT⁷ = (québ.) À VEILLE⁷, EN TRAIN⁷ = (québ.) APRÈS, etc.], ce qui n'est pas typique des grammèmes puisque cette variété suppose une certaine flexibilité en diachronie.

52. Les états, au sens de Vendler (1957).

53. Attention! D'autres acceptions de VIVRE sont compatibles avec ces signes [*Je viens de vivre une expérience inoubliable*].

5.7 Synthèse

Les faits découpent le temps réel en trois phases auxquelles on peut faire référence grâce aux auxiliaires AVOIR_{acc} et ALLER ainsi qu'à la locution prépositive 「EN TRAIN」. On peut également faire référence à l'environnement immédiat des bornes d'un fait grâce à 「SUR LE POINT」 et VENIR. Tous ces marqueurs dénotent des périodes.

Ces marqueurs de phase ne forment toutefois pas une catégorie flexionnelle au sens où nous l'entendons. D'une part, les marqueurs de phase ne sont pas mutuellement exclusifs :

(119) *Je vais avoir fini dans trois jours.*

(120) *Si tu arrives à l'heure du souper, ils vont être en train de manger!*

D'autre part, il faut remarquer que les marqueurs de phase ne se comportent pas comme des grammèmes puisqu'ils forment de nouvelles bases, qui peuvent elles-mêmes être modifiées sémantiquement par des grammies ou d'autres signes. D'ailleurs, les marqueurs de phase, comme tous les verbes, doivent porter des grammèmes de finitude, de mode et, le cas échéant, de temps. C'est ainsi que sont construits le futur périphrastique [*il va <allait> pleuvoir*], le passé récent [*(qu')(il) vient <venait/ vienne/venant/venir> de pleuvoir*], le passé composé à **valeur d'accompli** [*il a plu*], le plus-que-parfait à **valeur d'accompli** [*il avait plu*], les surcomposés [*il a <avait> eu plu*], le futur antérieur [*il aura plu*], le conditionnel passé [*il aurait plu*], le subjonctif passé [*qu'il ait plu*], le participe présent composé [*ayant plu*], l'infinitif passé [*avoir plu*], l'impératif passé [*aie fini!*], le gérondif passé [*en ayant fini*], ainsi que toutes les combinaisons de ces formes [*il allait avoir plu*, etc.]. Toutes ces formes portent un sens aspectuel (malgré ce que laisse croire la terminologie traditionnelle, qui a tendance à appeler «passé» tout ce qui contient l'auxiliaire AVOIR). Celles qui relèvent de l'indicatif portent, en plus de ce sens aspectuel, des indications temporelles, exactement de la même façon que les verbes simples correspondants, ce qu'on peut représenter schématiquement comme suit :

- 'pleut' = 'pleuvoir' \oplus 'g₁'
'a plu' = 'pleuvoir' \oplus 'avoir_{acc}' \oplus 'g₁'
- 'pleuvra' = 'pleuvoir' \oplus 'g₂'
'aura plu' = 'pleuvoir' \oplus 'avoir_{acc}' \oplus 'g₂'
- 'pleuvait' = 'pleuvoir' \oplus 'g₃'
'avait plu' = 'pleuvoir' \oplus 'avoir_{acc}' \oplus 'g₃'
- etc.

Ce qui nous intéresse maintenant, et qui sera l'objet de la prochaine section, est d'identifier les grammèmes qui permettent d'exprimer la composante temporelle des sens grammaticaux que nous avons représentés ici par la variable 'g_i'. Pour ce faire, il faut laisser de côté toutes les formes verbales dont on peut rendre compte par un (ou plusieurs) marqueurs de phase aspectuelle, et ne considérer que celles qui ne portent que des sens temporels.

6 Le temps

Nous avons jusqu'ici rendu compte des marqueurs de finitude, de mode et de phase. Nous sommes donc maintenant en mesure d'isoler les formes verbales qui portent des sens strictement temporels, à savoir :

- 1) Le présent [*Il regarde la télé*].
- 2) Le passé composé **quand il marque l'antériorité** [*Tout à coup, il lui a tendu la main*].
- 3) Le passé simple [*Il lui tendit la main*].
- 4) Le futur simple [*Quand les visiteurs arriveront, remettez-leur cette lettre*].
- 5) L'imparfait [*En 1998, elle travaillait à la banque*].
- 6) Le plus-que-parfait **quand il marque l'antériorité** [*Tout à coup, elle lui avait pris la main*].
- 7) Le conditionnel⁵⁴ **à valeur temporelle** [*Ils ont dit qu'ils viendraient*].

54. Rappelons que, sauf indication contraire, nous utilisons la terminologie traditionnelle pour nommer les formes verbales. Ce «conditionnel» n'exprime pas un sens modal.

Nous considérons les différences sémantiques entre toutes ces formes comme étant d'ordre temporel, et ce sont ces différences que nous tenterons d'expliquer ici. Rappelons que dans le présent chapitre nous accordons une attention particulière au sens de base de chaque grammème (la notion de sens de base d'un grammème, ainsi que les critères qui permettent de le déterminer, sont discutés au Ch. IV, § 4, p. 53). Les temps du français expriment presque tous plusieurs sens⁵⁵ et il vaut mieux s'en tenir dans un premier temps à un seul de ceux-ci (nous avons justifié cette approche au Ch. IV, § 3, p. 52).

6.1 Passé simple vs imparfait

Le système temporel du français présente la particularité de ne pas être symétrique puisque les temps du passé sont plus nombreux que ceux dénotant un présent ou un futur. Le trio passé simple ~ passé composé ~ imparfait n'a pas fini de faire couler de l'encre. Nous avons déjà démontré (§ 4.5, p. 158) que le passé composé correspond en fait à deux signes, AVOIR_{ant} et AVOIR_{acc}, ce dernier ayant déjà été traité plus haut (§ 5.2, p. 166). Dans ce qui suit, nous ne nous intéresserons donc qu'au passé composé à valeur temporelle, c'est-à-dire celui construit avec AVOIR_{ant}. Puisqu'on a tendance dans la littérature à confondre les deux acceptions de AVOIR, cette construction à valeur temporelle n'est pas systématiquement prise en compte, et c'est surtout l'opposition passé simple ~ imparfait qui a retenu l'attention. Afin d'alléger le texte, et en harmonie avec la tradition, nous parlerons dans les pages qui suivent de l'opposition «passé simple ~ imparfait», mais tout ce que nous disons à propos du passé simple s'applique également à AVOIR_{ant}. Nous considérons que AVOIR_{ant} et le passé simple sont de parfaits synonymes, si ce n'est la différence de registre (cf. § 6.3, p. 195).

Voyons d'abord quelques exemples (que nous empruntons au *Bescherelle* 1998 : § 144) qui illustrent l'opposition en question. Dans la phrase (121) ci-dessous, on comprend que les trois actions ont lieu simultanément (peut-être de façon intermit-

55. Considérons par exemple le présent dans *Son avion atterrit dans une heure* et dans *En 1287 l'abbé Lemieux se rend à Jérusalem*, qui situent les faits, respectivement, dans le futur et dans le passé.

tente sur un certain laps de temps), alors que (122) indique plutôt que les trois actions se succèdent.

(121) *Elle dansait, sautait et chantait.*

(122) *Elle dansa, sauta et chanta.*

Dans la phrase (123), le référent du verbe au passé simple est présenté comme s'insérant entièrement à l'intérieur du laps de temps correspondant au fait dénoté par le verbe à l'imparfait.

(123) *L'avion volait à haute altitude quand l'incident survint.*

Il y a donc quelque chose dans le sens de l'imparfait et du passé simple qui fait que le premier ne présente pas les faits comme des points dans le récit, alors que le second, oui. La question qui nous intéresse présentement est de savoir ce qui les distingue exactement, de façon à rendre compte des faits observés en (121)–(123), et comment cette opposition doit se représenter en termes de catégories flexionnelles. Bien entendu, nous ne considérons ici que le sens de base des formes en question. Il ne sera pas question dans cette section, par exemple, de l'imparfait «narratif», de l'imparfait «forain», etc. (nous en parlerons plus loin, § 6.4.4, p. 213).

6.1.1 *L'hypothèse de la durée*

Une des premières hypothèses proposées pour expliquer la différence entre le passé simple et l'imparfait a été la durée :

L'imparfait exprime les faits, même nouveaux, qui sont conçus comme ayant une certaine durée intérieure, le passé simple exprime les faits nouveaux conçus comme ayant un aspect nettement ponctuel; l'aspect reprend tous ses droits et devient plus que jamais décisif dans le choix des formes verbales.

Imbs (1960: p. 92)

Effectivement, le passé simple semble marquer des faits ponctuels, alors que l'imparfait semble au contraire marquer des faits ayant une certaine durée, comme dans la paire de phrases suivantes (nous empruntons ces exemples à Molendijk 1990: p. 21):

(124) *Jean me trouva. Je sortis.*

(125) *Jean me trouva. Je sortais.*

Il y a cependant des contre-exemples. Les plus évidents sont ceux où intervient un circonstanciel de temps indiquant une longue période, comme en (126):

(126) *Il dédia sa vie à combattre le crime.*

Molendijk (1990) est toutefois d'avis que ce genre de phrase ne constitue pas un contre-exemple réel puisque l'explicitation d'un intervalle ne présente pas forcément le fait comme «duratif»:

Il nous semble incorrect de la [= l'hypothèse d'une opposition ponctuel ~ duratif] rejeter en disant, par exemple, que le PS est compatible avec des indications de durée (cf. les observations faites par H. Weinrich (1971: 109) à propos du type 'ils s'injurièrent une heure durant'). Cette critique est trop facile. Si on associe la notion de durée à celle de continuation, on peut très bien défendre l'idée que le type 'ils s'injurièrent une heure durant' n'est pas duratif: cette phrase ne présente pas le fait qu'elle mentionne dans sa continuation, mais comme parvenant à son terme. Il y a donc un sens d'après lequel elle n'est pas durative.

Molendijk (1990: p. 20)

Il propose plutôt des exemples où un passé simple fait référence à un intervalle de temps plus grand qu'un imparfait avec lequel il peut commuter, comme dans le paragraphe suivant:

(127) *Tantôt, la chatte grimpait dans de grands arbres et faisait peur à tous les oiseaux de la montagne. Tantôt, elle s'étendait sur une roche et se faisait chauffer par le soleil. Elle s'amusa beaucoup ce jour-là.*

Molendijk (1990: p. 23)

Dans cet exemple, on comprend que la chatte s'est amusée toute la journée. Par contre, si on remplace *s'amusa* par *s'amusait*, on peut comprendre plutôt que la chatte s'est amusée durant un certain temps ce jour-là.

En raison de ces problèmes, nous croyons qu'il faut chercher ailleurs ce qui distingue le sens du passé simple de celui de l'imparfait. Voyons d'autres pistes qui ont été proposées.

6.1.2 *L'hypothèse de la perfectivité*

On a souvent utilisé l'«aspect» pour expliquer la différence sémantique entre le passé simple, qu'on dit porter l'aspect perfectif, et l'imparfait, qu'on dit imperfectif⁵⁶. Le terme d'«aspect» a été utilisé à diverses sauces, de sorte que les deux précisions terminologiques suivantes sont nécessaires.

Premièrement, nous n'utilisons jamais le terme «aspect» pour désigner des signifiants (par exemple, «aspect composé» — ou «aspect extensif» chez les guillaumiens — pour désigner les formes composées).

Deuxièmement, il ne faut pas confondre l'aspect grammatical (celui qui nous intéresse principalement) et l'**Aktionsart** (aussi appelé «mode d'action»). Ce dernier est une propriété intrinsèque du sens des unités lexicales (ou plus précisément, des groupes verbaux), alors que l'aspect grammatical fait partie du système grammatical de la langue. Dans le cadre de notre travail, nous utilisons la classification des aspects lexicaux de Vendler (1957)⁵⁷ telle que redéfinie par Vetters (1996), qui classe les syntagmes verbaux⁵⁸ en quatre catégories⁵⁹:

- Les **états** dénotent des faits qui ne sont ni dynamiques, ni téliques ni ponctuels. Par exemple: *Il s'appelle Paul.*
- Les **activités** dénotent des faits dynamiques mais atéliques qui ne sont pas ponctuels. Par exemple: *Elle se promène.*
- Les **accomplissements** dénotent des faits dynamiques et téliques, mais non ponctuels. Par exemple: *Demain nous irons à Chengdu.*
- Les **achèvements** (ou «réalisations instantanées») dénotent des faits dynamiques, téliques et ponctuels. Par exemple: *Il atteint le sommet.*

La définition classique de l'**aspect grammatical** est celle qu'on retrouve chez Comrie (1976: p. 3): «aspects are different ways of viewing the internal tem-

56. Voir notamment Comrie (1976) ou Gosselin (1996).

57. Garey (1957) et Comrie (1976) ont aussi proposé des classifications similaires.

58. Pour Vendler, c'est tout le syntagme qui doit être considéré, et non le verbe seul.

Comparer, par exemple *Il mange* (qui n'est pas télique) et *Il mange sa soupe* (qui l'est).

59. Nous reviendrons sur ces catégories au Chapitre VIII (§ 1.2, p. 285).

poral constituency of a situation». Le concept central dans cette définition est celui de structure temporelle *interne*. Ce qui distingue l'aspect du temps grammatical est que le premier indique comment se déroule le procès, alors que le second situe ce même fait dans le temps réel. L'aspect n'est pas, contrairement au temps grammatical, déictique: il ne met en jeu rien d'autre que le fait lui-même, dont on considère la structure temporelle interne, alors que le temps grammatical implique un point temporel externe au fait, par rapport auquel est situé ce dernier.

L'opposition aspectuelle la mieux documentée est probablement celle entre le perfectif et l'imperfectif⁶⁰. L'aspect perfectif présente le fait comme un tout, sans égard à sa constitution temporelle interne (le début, le déroulement et la fin ne sont pas distincts). L'imperfectif, quant à lui, fait référence à la constitution temporelle interne du fait, mais pas à son début et à sa fin. Cette opposition aspectuelle semble se retrouver en français dans l'opposition passé simple⁶¹ ~ imparfait. Veters (1996: pp. 114–118) relève plusieurs indices de la perfectivité du passé simple et de l'imperfectivité de l'imparfait. Les indices de perfectivité sont les suivants (tous les exemples sont de lui⁶²):

- L'explicitation de la fin de l'action: *Le 5 juin 1989, Jules attendit <* attendait> jusqu'à 5 heures.*
- L'indication de l'espace temporel du début jusqu'à la fin: *Le 5 juin 1989, Jules étudia <* étudiait> du matin jusqu'au soir.*
- L'indication de la durée totale: *Le 5 juin 1989, Jules rentra <* rentrait> chez lui en 50 minutes.*
- La répétition définie d'événements particuliers: *Le 5 juin 1989, Jules sonna <* sonnait> trois fois à la porte.*

60. Voir notamment Mazon (1914) et Comrie (1976).

61. Rappelons que dans la présente section, tout ce que nous disons à propos du passé simple est aussi valide pour le passé composé construit avec AVOIR_{ant}.

62. Veters laisse volontairement de côté l'imparfait de narration dans ses exemples. Cela est tout à fait compatible avec notre méthodologie, qui consiste à ne considérer dans un premier temps que le sens de base des grammèmes.

- L'indication de la réalisation instantanée: *Tout à coup, Jean tomba.*
- Les changements d'état ou de position: *Quelques secondes plus tard, Luc fut sous le chapiteau.*

Au sujet de ce dernier indice, il dit:

Les états sont en principe caractérisés par le fait qu'ils ne changent pas. S'il y a quand même un changement, cette transition entre les deux états est considérée comme un événement perfectif (et ponctuel) et permet l'emploi de l'aspectif⁶³ de perfectivité.

Vetters (1996: p. 116)

Les indices d'imperfectivité, quant à eux, sont (nous empruntons encore à Vetters):

- Le décor: *Nous étions <* fûmes> à l'étude quand le proviseur entra* (Flaubert, *Madame Bovary*, p. 37).
- L'événement interrompu: *Il se noyait <* noya> quand l'agent le sauva en le retirant de l'eau.*
- Discordance entre un événement bref et une période de temps étendue: *À cette époque, Jean prenait <* prit> son café avec moi.*
- La répétition non spécifiée: *À cette époque, Paul tombait souvent ~ * À cette époque, Paul tombait.*

Il est à noter que tant les indices de perfectivité que d'imperfectivité sont sans effet avec le conditionnel et le futur simple. En remplaçant les passés simples et les imparfaits dans les exemples de Vetters par ces formes, on obtient toujours des phrases grammaticales. Par contre, le présent a la même distribution que l'imparfait; il est aussi imperfectif.

Ces faits suggèrent fortement que le passé simple et l'imparfait se distinguent essentiellement par l'aspect, et c'est l'hypothèse qui prédomine aujourd'hui. Si on l'accepte, on doit alors se poser la question suivante: l'aspect forme-t-il une catégorie flexionnelle à part entière qui se combine à la catégorie de temps, ou est-il une simple composante sémantique des grammèmes de temps? Pour la quasi-tota-

63. Un «aspectif», chez Vetters, est un marqueur d'aspect.

lité des linguistes qui se sont intéressés de près à cette opposition aspectuelle dans les temps du passé en français, le question n'a pas vraiment de pertinence étant donné leur cadre méthodologique. Pour nous, elle est essentielle. Elle l'était également pour Martinet (1979), qui disait à propos de l'opposition entre les «monèmes» *passé* et *prétérit* (c'est-à-dire entre l'imparfait et le passé simple):

Le monème « passé » place le procès dans le flux du temps révolu au moment de l'acte de parole. [...] Par opposition au « prétérit », cerné, il se définit comme 'non limité'. On pourrait sans doute faire valoir que l'opposition entre passé et prétérit n'est pas proprement de temps, puisque l'un et l'autre se placent dans le même temps révolu par rapport à l'instant de l'acte de parole. Si les deux monèmes passé et prétérit sont rangés ici dans une même classe, c'est qu'ils sont mutuellement exclusifs.

Martinet (1979: p. 106)

Pour lui, l'aspect est donc une partie intégrante du sens des grammèmes de temps. Si on accepte cette hypothèse, on a alors une seule catégorie flexionnelle de temps, dont certains grammèmes expriment des sens strictement temporels (le futur simple et le conditionnel), alors que d'autres (le présent, l'imparfait, le passé composé construit avec AVOIR_{ant} et le passé simple) expriment des sens à la fois temporels et aspectuels.

Aspectuels		Inaspectuels
Perfectifs	Imperfectifs	
<i>il mangea</i> <i>il a mangé</i> <i>il avait mangé</i>	<i>il mange</i> <i>il mangeait</i>	<i>il mangera</i> <i>il mangerait</i>

Tableau XVII — L'hypothèse des temps à valeur aspectuelle intrinsèque

Le manque d'homogénéité sémantique et la non-symétrie des grammèmes exprimant le futur (sans aspect)⁶⁴ ou le passé (avec aspect) sont légèrement gênants, mais pas assez pour rejeter cette hypothèse.

64. Vet (1980) soutient que le futur simple est perfectif en se basant notamment sur les verbes d'état, qui semblent prendre une valeur inchoative au futur: *Jeanne saura la réponse (demain)* (Vet 1980: p. 82). Cette interprétation n'est toutefois pas intrinsèque au futur, comme le montre la phrase *Je le savais déjà hier, je le sais encore, et je le saurai toujours demain.*

L'hypothèse contraire consiste à stipuler deux catégories flexionnelles: une catégorie de temps et une d'aspect. La catégorie de temps comprendrait un grammème **passé**, auquel se combineraient les grammèmes d'aspect, **perfectif** et **imperfectif**. L'imparfait serait alors le marqueur de deux grammèmes, **passé** \oplus **imperfectif**, tandis que le passé simple et AVOIR_{ant} seraient les marqueurs de la combinaison **passé** \oplus **perfectif**. Cette solution n'a pas les inconvénients de la première; par contre, elle a le défaut d'ajouter une catégorie flexionnelle au modèle. Il faut se demander si elle est vraiment nécessaire.

En fait, d'autres linguistes ont carrément remis en question l'opposition aspectuelle comme principal élément distinctif entre le passé simple et l'imparfait. En particulier, Molendijk (1990) réfute cette hypothèse en montrant des exemples où l'imparfait a une valeur perfective, c'est-à-dire qu'il présente le fait dénoté dans son intégrité. Les deux exemples suivants (de Molendijk 1990: pp. 18–20) nous suffiront à résumer sa principale objection:

(128) *Hélène était la fille du roi de Pologne.*

(129)**Hélène était la fille du roi de Pologne lorsque, soudain ...*

En (128), le fait dénoté par *était la fille* est une propriété inhérente qui a cours tout au long de la vie d'Hélène; il a donc un contour temporel bien défini, ce qui en ferait selon Molendijk un perfectif. L'idée est encore renforcée par l'agrammaticalité de (129), qui montre que la structure temporelle interne du fait dénoté par *était la fille* n'est pas accessible. Ainsi, ce n'est pas l'aspect qui distingue le passé simple et l'imparfait.

6.1.3 *L'hypothèse de la progression narrative*

Kamp & Rohrer (1983), sans explicitement rejeter l'hypothèse aspectuelle pour expliquer la différence entre le passé simple et l'imparfait, proposent une alternative. Selon eux, il ne suffit pas de dire que les temps du passé situent les faits avant le moment d'énonciation puisqu'ils les situent également par rapport aux autres faits passés, comme dans le paragraphe suivant:

(130) *Le professeur entra, s'installa au tableau et donna son cours de lexicologie. Il parlait très fort.*

Les faits dénotés par les verbes *entra, s'installa* et *donna* ne se produisent pas dans n'importe quel ordre : ils sont présentés comme successifs. Le référent de *parlait*, quant à lui, couvre toute la durée du cours. Kamp & Rohrer (1983) proposent donc de définir le passé simple et l'imparfait en termes d'instructions discursives comme suit :

Le passé simple

- 1) Le passé simple introduit un fait *e* qui précède le moment d'énonciation.
- 2) *e* suit le dernier événement déjà introduit.

L'imparfait

- 1) L'imparfait introduit un état *s* qui précède le moment d'énonciation.
- 2) *s* contient le dernier événement *e*.

La différence entre les deux réside donc dans le fait que le passé simple fait progresser le récit, c'est-à-dire qu'il déplace le point de référence temporelle, alors que l'imparfait ne produit pas cet effet. Toutefois, il y a des contre-exemples. Kamp & Rohrer (1983 : p. 260) eux-mêmes relèvent les trois cas suivants :

- (131) *Marie chanta et Pierre l'accompagna au piano.*
- (132) *L'année dernière Jean escalada le Cervin. Le premier jour il monta jusqu'à la cabane H. Il y passa la nuit. Ensuite il attaqua la face nord. Douze heures plus tard il arriva au sommet.*
- (133) *L'été de cette année-là vit plusieurs changements dans la vie de nos héros. François épousa Adèle, Jean-Louis partit pour le Brésil et Paul s'acheta une maison à la campagne.*

En (131), les deux faits sont simultanés. En (132), l'escalade de Jean couvre une période qui englobe tous les autres faits présentés dans le paragraphe. Quant au paragraphe (133), il n'est pas possible de savoir dans quel ordre se sont produits les faits mentionnés dans la deuxième phrase. Ces contre-exemples ont poussé Kamp & Rohrer (1983) à reformuler leur définition du passé simple ainsi :

Le passé simple

- 1) Le passé simple introduit un fait *e* qui précède le moment d'énonciation.
- 2) *e* ne peut pas complètement précéder l'événement ou le temps qui sert de point de référence.

Malgré tout, cette définition se heurte au contre-exemple suivant, noté par Vetters (1996) en reprenant un exemple de Molendijk, où le second passé simple dénote un événement antérieur à celui que dénote le premier :

- (134) *En 1982, il s'installa dans la ville même où Charles Martel arrêta les Arabes.*

Molendijk (1990: p. 251)

Des contre-exemples ont aussi été proposés pour invalider la définition de l'imparfait de Kamp & Rohrer (1983). Un de ces contre-exemples est l'imparfait dit «narratif» [*Le lendemain matin, l'armée entra dans la ville pour mater la rébellion*]. Cet imparfait a la valeur d'un passé simple. Nous croyons toutefois que, puisqu'il ne s'agit pas du sens de base de l'imparfait, il ne peut pas constituer un véritable contre-exemple. Cependant, l'exemple suivant de Molendijk (1990) est construit sur l'acception de base de l'imparfait et contredit la définition de Kamp & Rohrer (1983) puisque cet imparfait ne reprend pas comme point de référence le fait introduit par le dernier passé simple (exemple original de Molendijk 1990, que nous avons modifié):

- (135) *Pierre vira à droite et passa à tout allure devant un policier qui l'arrêta et lui donna une contravention. Il roulait beaucoup trop vite.*

6.1.4 *L'approche reichenbachienne*

En réponse notamment aux problèmes liés aux hypothèses que nous venons de présenter, Molendijk (1990) propose une solution alternative basée sur la représentation des temps de l'anglais de Reichenbach (1947), que nous allons d'abord résumer.

La thèse principale de Reichenbach est que le sens des temps ne peut pas toujours se représenter seulement en termes de relations temporelles entre le fait

dénoté par le verbe et le moment d'énonciation, mais qu'intervient aussi un troisième point temporel, le point de référence. Les faits, plutôt que d'être directement situés par rapport au moment d'énonciation, le sont par rapport au point de référence, qui lui-même l'est par rapport au moment d'énonciation. En représentant le moment où se déroule le fait en question par *E* (pour *event*), le moment d'énonciation par *S* (pour *speech*) et le point de référence par *R* (pour *reference*), on peut voir qu'il y a trois relations possibles entre *E* et *S*, ainsi qu'entre *R* et *S*, que Reichenbach nomme ainsi :

Relation E/R	Nom	Relation R/S	Nom
$E < R$	Antérieur	$R < S$	Passé
$E = R$	Simple	$R = S$	Présent
$E > R$	Postérieur	$R > S$	Futur

Tableau XVIII — Les relations simples E/R et R/S selon Reichenbach

La relation entre *S* et *E* n'est pas pertinente dans son modèle. Les deux autres relations se combinent, ce qui donne en théorie les neuf relations temporelles complexes suivantes entre *E*, *R* et *S*⁶⁵ :

	Passé ($R < S$)	Présent ($R = S$)	Futur ($R > S$)
Antérieur ($E < R$)	<i>He had eaten</i>	<i>He has eaten</i>	<i>He will have eaten</i>
Simple ($E = R$)	<i>He ate</i>	<i>He eats</i>	<i>He will eat</i>
Postérieur ($E > R$)	<i>He would eat</i>	<i>He will eat</i>	—

Tableau XIX — Les relations complexes E/R/S selon Reichenbach

Au moins trois remarques s'imposent ici :

- 1) La relation entre *S* et *E* peut généralement être déduite des relations *E/R* et *R/S*. Par exemple, pour le «passé antérieur» : $E < R \ \& \ R < S \rightarrow E < S$. Les

65. Ces exemples sont de nous mais sont proches de ceux de Reichenbach, qui utilisait l'auxiliaire *shall* au lieu de *will*.

deux seules exceptions sont le «futur antérieur» et le «passé postérieur». Dans les deux cas, il n'est pas possible de déterminer où se situe le référent du verbe par rapport au moment d'énonciation, ce qui correspond bien aux faits linguistiques.

- 2) Le modèle prévoit une case, celle du «futur postérieur», qui n'a pas d'équivalent en anglais. Elle correspond, selon Reichenbach, au participe futur du latin.
- 3) La forme *will eat* reçoit deux interprétations différentes: «futur simple» et «présent postérieur», ce qui se justifie d'après l'auteur par le fait qu'on peut dire *Now I shall go*, qui s'interpréterait comme un «présent postérieur» ($R = S \ \& \ E > R$), et *I shall go tomorrow*, qui serait plutôt à interpréter comme un «futur simple» ($R > S \ \& \ E = R$).

Le modèle de Reichenbach est basé sur l'anglais mais prétend à l'universalité. On devrait donc pouvoir l'utiliser pour le français. Or, là où l'anglais n'a que le passé simple, le français possède l'imparfait et le passé simple (ou son synonyme, AVOIR_{ant}). Reichenbach attribue à ces deux formes la même formule ($R < S \ \& \ E = R$), mais représente l'imparfait comme un temps «étendu» et le passé simple comme un temps «non étendu», ce qui revient au même qu'une opposition duratif ~ ponctuel (§ 6.1.1, p. 179) ou imperfectif ~ perfectif (§ 6.1.2, p. 181), pour lesquelles nous avons déjà exposé nos réserves. Le modèle de Reichenbach, tel quel, ne permet donc pas de représenter de façon satisfaisante l'opposition passé simple ~ imparfait. Il semble cependant constituer un point de départ intéressant pour notre travail; nous allons donc en discuter un peu plus en détail.

Vetters (1996: pp. 18–27) a si bien synthétisé les diverses critiques qui ont été faites du modèle reichenbachien que nous ne pouvons que le paraphraser en le résumant et renvoyer à ce texte le lecteur intéressé par une discussion plus en profondeur. Ces critiques s'articulent essentiellement autour de quatre axes.

Premièrement, les formes progressives de l'anglais ne sont pas traitées de façon satisfaisante dans ce modèle. Le problème réside dans le fait que la différence entre

les formes progressives et non-progressives est essentiellement aspectuelle, alors que le modèle de Reichenbach ne peut rendre compte que d'informations temporelles. Ce problème semble à première vue particulièrement important quand on tente de décrire l'opposition passé simple ~ imparfait en français, qui est, pour plusieurs linguistes, d'ordre aspectuel. Nous croyons cependant qu'il ne faut pas confondre temps et aspect, et qu'un modèle du temps grammatical comme celui de Reichenbach n'a pas à résoudre également le problème de l'aspect.

Deuxièmement, les neuf formes prévues par la combinaison des deux séries de trois relations temporelles élémentaires ne correspond pas bien aux faits linguistiques. Pour commencer, la case vide du «futur postérieur» ne semble pas trouver d'écho dans les langues. Le participe futur du latin proposé par Reichenbach n'est pas satisfaisant puisqu'il ne renvoie pas directement à un fait, mais plutôt à sa préparation. La forme qui est normalement utilisée pour exprimer le «futur postérieur» est le futur simple (*He will say that he will eat it the next day*). Il faudrait donc en principe ajouter cette forme à la case vide du Tableau XIX. Or, cela ne fait qu'ajouter un problème. Cette forme se trouve déjà dans deux cases du tableau, ce qui n'est pas justifié d'un point de vue linguistique puisqu'on ne peut pas démontrer qu'il y a deux futurs en anglais. Reichenbach propose une solution basée sur l'opposition entre le futur périphrastique (*il va manger*) et le futur simple (*il mangera*) en français: le premier serait un «présent postérieur» alors que le second serait un «futur simple» selon sa terminologie. Cependant, la solution ne résout pas vraiment le problème puisque le futur simple continue d'occuper deux cases (celles du «futur simple» et du «futur postérieur») sans justification solide. Par ailleurs, si on fait entrer le futur périphrastique dans ce modèle, il faut alors considérer aussi la forme *allait manger*, pour laquelle il n'y a pas de case libre (nous supposons qu'il s'agirait d'un «passé postérieur», mais cette case est déjà occupée par le conditionnel en français — il faudrait alors démontrer que ces deux formes sont de vrais synonymes).

Troisièmement, alors que le modèle de Reichenbach se veut universel, il ne résiste pas à l'épreuve des faits quand on tente de l'appliquer tel quel à des langues

qui ont, par exemple, des grammèmes exprimant la distance dans le temps ou qui possèdent une catégorie de temps relatif (dont les grammèmes expriment des relations temporelles n'impliquant pas le moment d'énonciation). Il serait alors nécessaire d'ajouter au modèle un axe supplémentaire.

Quatrièmement, Reichenbach ne donne pas de définition précise du point de référence. En outre, les adverbes de temps, d'après lui, spécifient toujours *R*, mais il existe des contre-exemples, dont celui-ci de Vetters (1996: p. 23): *Jean était arrivé à 5 heures*. Cette phrase peut s'interpréter comme indiquant l'heure de l'arrivée de Jean, auquel cas l'adverbe de temps spécifie *E* plutôt que *R*.

Molendijk (1990) s'est particulièrement intéressé à la définition de *R* dans le système reichenbachien. Sa définition prend quelques pages, nous n'allons donc pas l'exposer ici; il nous importe seulement de remarquer qu'il cherche essentiellement à répondre à la question suivante: «Quel est le fait (moment) auquel il faut identifier *R*, après l'apparition d'une phrase *P* appartenant à un texte donné?» (Molendijk 1990: p. 127). En d'autres termes, il ne s'intéresse qu'à l'analyse. C'est là une caractéristique commune à la très grande majorité des modèles qui ont été proposés pour rendre compte du temps (notamment dans un cadre discursif / narratif). Or, nous croyons que c'est faire fausse route que de s'attacher à décrire le processus d'analyse avant tout. La langue ne sert pas seulement à comprendre, mais aussi (et surtout) à s'exprimer. Il nous semble peu probable que le locuteur, quand il s'exprime, suive des règles qui lui imposent tel ou tel autre point de repère temporel. Il sait par rapport à quoi il situe les faits, puisque c'est lui qui décide. Il doit seulement s'arranger pour que son interlocuteur puisse décoder son message. Mais ce que des modèles comme ceux de Molendijk (1990) ou de Kamp & Rohrer (1983) décrivent, ce sont les *processus cognitifs* qui permettent au récepteur de décoder un message. Ces processus doivent bien sûr être décrits, mais ce n'est pas notre propos de le faire. Nous nous intéressons uniquement à décrire les *signes* de la langue, ce qui ne saurait se faire en termes de règles calculant la position d'un point de référence.

À notre avis, ce qui constitue le principal problème du modèle de Reichenbach, c'est qu'il cherche à décrire dans un même système des signes temporels et des signes aspectuels, ce qui le conduit à une définition vague de *R*. À la seconde ligne du Tableau XIX (p. 188), on peut voir que le passé simple, le présent et le futur simple sont représentés comme encodant tous l'information $E=R$, c'est-à-dire que le moment où se produit le fait en question sert lui-même de point de référence. Or, nous ne voyons pas en quoi la phrase *He ate beans*, par exemple, indique que le fait de manger est situé par rapport au moment du repas, et non par rapport au moment d'énonciation. Il nous semble qu'il s'agit là d'une conséquence d'avoir voulu intégrer au système les formes construites avec HAVE, qui correspondent en anglais à celles construites avec AVOIR_{acc} en français, c'est-à-dire le marqueur de l'aspect accompli. On se demande alors pourquoi ne pas avoir intégré également les formes en 'BE GOING'. Il est clair que ces formes ne trouvent pas leur place dans ce système, où il n'est pas possible de cerner la différence sémantique entre les auxiliaires WILL et 'BE GOING'.

Environ 80 ans plus tôt, Te Winkel (1866)⁶⁶ avait proposé pour les langues germaniques un système similaire, mais construit selon trois axes, plutôt que les deux de Reichenbach. Un premier axe indique le point à partir duquel se conçoit la pensée: soit par rapport au moment d'énonciation, soit par rapport à un temps dans le passé. Un deuxième axe situe les faits comme synchrones ou postérieurs au point de référence. Les formes du parfait sont représentées sur un troisième axe qui croise les deux premiers et qui indique si le fait est complété ou non. Le Tableau XX ci-dessous résume ce modèle, qui constitue à notre avis un meilleur point de départ que celui de Reichenbach puisqu'il sépare bien le temps de l'aspect.

Nous verrons qu'un modèle similaire peut être construit pour le français et qu'il est possible de rendre compte de l'opposition passé simple ~ imparfait sans avoir recours à une opposition aspectuelle qui ne s'appliquerait qu'aux temps du passé.

66. C'est Verkuyl (2005) qui a pour ainsi dire ressuscité Te Winkel.

		Passé	Présent
Non complété	Synchrone	<i>He ate</i>	<i>He eats</i>
	Postérieur	<i>He would eat</i>	<i>He will eat</i>
Complété	Synchrone	<i>He had eaten</i>	<i>He has eaten</i>
	Postérieur	<i>He would have eaten</i>	<i>He will have eaten</i>

Tableau XX — Le modèle de Te Winkel

Mais d'abord, nous discuterons une autre paire de temps qui a attiré l'attention: le présent et l'imparfait.

6.2 Présent vs imparfait

Le présent et l'imparfait présentent d'importantes similarités de sens; ce fait est d'ailleurs bien connu depuis longtemps. Le Goffic (1986b: p. 62) nous rappelle que les Stoïciens l'avaient déjà remarqué à propos du grec ancien, qui avait un présent et un imparfait semblables à ceux du français. Ces similarités doivent nous dire quelque chose à propos de l'organisation des grammèmes temporels en français.

Labeau (2002) a recensé les valeurs de l'imparfait dans cinq grammaires récentes. Nous avons fait la même chose avec les valeurs du présent dans le *Bescherelle* (1998: § 143), Grevisse (1993: § 850) et Riegel, Pellat & Rioul (1994: pp. 299–301). Presque toutes les valeurs du présent trouvent leur écho à l'imparfait, et vice-versa. Le Tableau XXI ci-dessous résume les différentes acceptions de ces deux formes.

	Présent	Imparfait
Actualité	<i>Il pleut.</i>	<i>Il pleuvait.</i>
Validité permanente	<i>Le soleil se couche à l'Ouest.</i>	<i>Il savait que le soleil se couchait à l'Ouest.</i>
Habitude	<i>Il mange quatre fois par jour.</i>	<i>Il mangeait quatre fois par jour à l'époque.</i>

Tableau XXI — Les sens du présent et de l'imparfait dans quelques grammaires de référence

VI. Un modèle descriptif de la conjugaison française

Passé récent	<i>Je rentre tout juste de Berlin</i>	<i>Je rentrais tout juste de Berlin</i>
Futur inéluctable	<i>Nous partons à 5h demain.</i>	<i>Nous partions à 5h le lendemain.</i>
Historique	<i>Le 2 octobre 1535, Cartier arrive à Hochelaga.</i>	<i>Le 2 octobre 1535, Cartier arrivait à Hochelaga.</i>
Conditionnel	<i>S'il pleut, on ira au cinéma.</i>	<i>Elle a dit que s'il pleuvait, on irait au cinéma.</i>
Injonctif	<i>On se calme!</i>	—
Atténuation	—	<i>Je voulais vous demander quelque chose.</i>
Irréel	—	<i>Si j'étais riche, je partirais en voyage autour du monde.</i>

Tableau XXI — Les sens du présent et de l'imparfait dans quelques grammaires de référence

Le parallèle entre les deux formes est flagrant. Pour les sept premières acceptions, le présent et l'imparfait se distinguent uniquement par l'ancrage temporel des faits dénotés: l'imparfait exprime par rapport à un repère passé ce que le présent exprime par rapport au moment d'énonciation. Ces faits suggèrent que le présent et l'imparfait ne sont pas indécomposables, mais qu'ils sont plutôt la composition de deux grammèmes: un qu'ils partagent (et dont les différentes acceptions correspondent aux emplois parallèles de ces deux formes) et un autre qui les distingue (en indiquant où se trouve le repère temporel).

Toutefois, le présent a aussi une acception qui ne trouve pas son équivalent à l'imparfait, et l'imparfait a deux acceptions qui lui sont propres. Le parallèle entre les deux formes n'est donc pas parfait. Cependant, nous verrons que ce manque apparent de symétrie ne permet pas de rejeter l'hypothèse que nous venons de formuler. En fait, un bref survol du sens de base de chacune des formes verbales à valeur temporelle nous confortera dans cette hypothèse.

6.3 Caractérisation sémantique des «formes temporelles» du français

Nous avons donné au début de cette section sur le temps (§ 6, p. 177) la liste des tiroirs purement temporels: le présent, le passé composé d'antériorité, le passé simple, le futur simple, l'imparfait, le plus-que-parfait d'antériorité et le conditionnel temporel. C'est uniquement par l'étude de ces formes que nous arriverons à comprendre le fonctionnement des grammèmes de temps en français, puisque les autres formes expriment des sens aspectuels ou modaux qui brouillent les pistes. Nous avons vu plus haut (Ch. VI, § 4.5, p. 158) qu'il existe deux auxiliaires **AVOIR**: un à valeur d'accompli, l'autre à valeur d'antériorité. Les formes du passé composé et du plus-que-parfait sont ambiguës, puisqu'elles peuvent exprimer les deux sens. Nous ne considérons ici que leur interprétation strictement temporelle. Les autres formes construites avec **AVOIR** ont une valeur d'accompli, et sont donc laissées de côté dans les pages qui suivent, au même titre que les formes construites avec les autres marqueurs de phase aspectuelle.

Dans les pages qui suivent, nous allons donner une approximation du sens temporel principal exprimé par chacun des sept tiroirs verbaux purement temporels, à l'aide de schémas qui montrent, sur un axe du temps, comment les faits sont situés par rapport au moment d'énonciation. Les caractérisations sémantiques que nous donnons ci-dessous pour chaque forme ne sont valables que pour l'acception en jeu dans les exemples que nous avons donnés plus haut (§ 6, p. 177), c'est-à-dire l'acception de base des grammèmes de temps impliqués dans ces formes. Nous devons anticiper un peu sur notre conception du temps grammatical, que nous ne pourrions justifier que dans les sections subséquentes. Dans nos figures, nous représentons le moment d'énonciation par le symbole T_0 et nous indiquons par T_1 le moment de référence par rapport auquel sont situés les faits lorsqu'il diffère de T_0 ⁶⁷.

Il faut garder à l'esprit que les éléments de la liste ci-dessus, que nous nous apprêtons à étudier un à un, sont des *formes* verbales, et non des grammèmes.

67. Notre T_0 correspond au S de Reichenbach (1947) (cf. § 6.1.4, p. 187). Nous verrons plus loin que notre T_1 , en revanche, diffère légèrement de son R (cf. § 6.4, p. 202).

Nous ne présupposons pas qu'à chacun de ces tiroirs corresponde un grammème. En fait, nous avons déjà émis l'hypothèse que le présent et l'imparfait puissent être décomposés en grammèmes plus simples (*cf.* § 6.2, p. 193). D'après notre modèle, aucune de ces formes n'est élémentaire. Toutes expriment plus d'un grammème. Il ne faut donc surtout pas croire que le présent, le conditionnel, le passé simple, etc., sont des éléments de catégories flexionnelles dans notre modèle. Nous ne cherchons ici qu'à donner une approximation du sens porté par ces formes, pour ensuite en tirer des conclusions sur la nature des grammèmes temporels qu'elles expriment (§ 6.4, p. 202).

6.3.1 *Le présent*

Parmi les nombreux sens du présent (*cf.* § 6.2, p. 193), celui qui constitue le sens de base (principalement en vertu du critère de l'interprétation spontanée) pose un fait comme contemporain au moment d'énonciation [*Il regarde le mur*]:

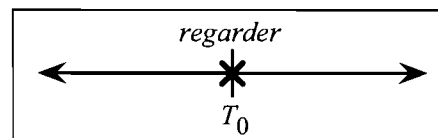


Figure 23 — *Caractérisation sémantique du présent*

Les fonctionnalistes, comme Martinet (1979) ou encore Touratier (1996), voient dans le présent une absence de temps (*cf.* Ch. V, § 3, p. 84). En fait, ils ne le considèrent pas comme un morphème car selon eux le présent dénote des sens tellement variés qu'il est plus simple de dire qu'il ne dénote rien et que tous ces effets de sens viennent du contexte. Cette position découle de l'approche non discrète qu'ils ont adoptée et néglige cependant le fait que, s'il est vrai que le présent peut effectivement porter un grand nombre de sens, il ne peut pas porter n'importe quel sens. Si on accepte que, comme les signes lexicaux, les signes grammaticaux sont des unités discrètes et qu'il peut y avoir des grammèmes polysémiques, alors le problème de la variété des sens du présent ne se pose pas.

6.3.2 Le passé simple et le passé composé

Le passé simple [*Il tendit la corde*] et le passé composé construit avec l'auxiliaire AVOIR_{ant} [*Tout à coup, il a tendu la corde*] ont exactement le même sens. Les deux situent les faits dans le passé par rapport au moment d'énonciation :

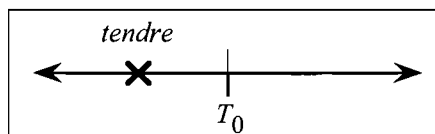


Figure 24 — Caractérisation sémantique du passé

Plusieurs linguistes voient entre le passé simple et le passé composé une différence aspectuelle⁶⁸. Cette idée provient, à notre avis, d'une confusion due à l'ambiguïté du passé composé, qui peut marquer soit l'accompli, soit l'antériorité, selon qu'il est construit avec AVOIR_{acc} ou AVOIR_{ant} (cf. § 4.5, p. 158). Il est vrai néanmoins que dans certains contextes le passé simple peut sembler ne pas commuter avec un passé composé. Par exemple :

(136) *Ayant fait le tour de la ville, je rentraï.*

(137)[?] *Ayant fait le tour de la ville, je suis rentré.*

Nous croyons toutefois que ce contraste de grammaticalité est dû à une question de registre de langue. Le passé simple n'est plus vraiment utilisé à l'oral, il appartient clairement à un registre différent, au style littéraire. Le passé composé, quant à lui, est plutôt d'un registre neutre. Le fait que la phrase (136) semble meilleure que la phrase (137) s'expliquerait donc par une plus grande concordance du registre de *ayant fait* (expression elle-même d'un style soutenu) et de celui du passé simple. Si on change *ayant fait* pour une construction plus neutre stylistiquement, le contraste de grammaticalité disparaît, et nous ne sentons pas de différence de sens entre les deux phrases :

(138) *Après avoir fait le tour de la ville, je rentraï.*

(139) *Après avoir fait le tour de la ville, je suis rentré.*

68. Par exemple, Imbs (1960), Comrie (1976), ou Gosselin (1996).

Donc, le fait qu'il existe deux formes concurrentes pour exprimer l'antériorité par rapport au moment d'énonciation n'est pas gênant, puisqu'il y a une différence de registre entre ces formes. Cette irrégularité s'explique par le fait que le remplacement du passé simple par le passé composé pour marquer l'antériorité est une évolution relativement récente du français, qui n'est pas encore achevée.

Nous ne connaissons pas d'autre acception au passé simple, et n'en avons trouvé mention dans aucune des grammaires de références consultées. Grevisse (1993: § 852) remarque que le passé simple peut exprimer l'antériorité par rapport à un autre passé [*Cette femme [= une actrice] à l'agonie [...] témoignait d'une grandeur que, devant la rampe, elle n'atteignit jamais*]. Il ne faut pas en conclure, dit-il, que le passé simple sert à marquer l'antériorité par rapport à d'autres faits, ce avec quoi nous sommes parfaitement d'accord. Il n'y a rien qui indique, en effet, que le passé simple situe les faits par rapport à d'autres faits (passés ou futurs), et non par rapport au moment d'énonciation.

Quant au passé composé, on lui attribue dans à peu près toutes les grammaires au moins deux sens, mais nous y voyons une confusion entre les auxiliaires AVOIR_{ant} et AVOIR_{acc} (cf. § 4.5, p. 158). Nous n'avons pas trouvé d'autre sens à AVOIR_{ant} que celui décrit ici.

6.3.3 *Le futur simple*

Dans son interprétation de base, le futur simple [*Quand les visiteurs arriveront, remettez-leur cette lettre*] situe (tout simplement) un fait dans le futur par rapport au moment d'énonciation :

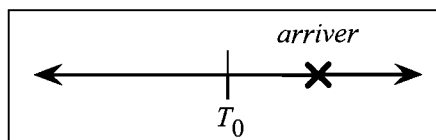


Figure 25 — Caractérisation sémantique du futur simple

Le futur simple peut également avoir un sens modal indiquant une supposition de la part du locuteur [*J'entends des pas. Ce sera sans doute Lucie*]. Cependant, il

ne peut s'agir du sens de base puisqu'il échoue au critère de l'interprétation spontanée. Le futur dit «historique» [*Le 15 mai 1940, l'état-major néerlandais capitulera sans condition*] est aussi d'un emploi trop limité pour constituer le sens de base. Enfin, le futur «d'atténuation» [*Je ne vous cacherai pas que j'ai eu peur*] est d'un emploi encore plus limité (il ne peut apparaître que dans une proposition principale, et à la première personne seulement).

6.3.4 L'imparfait

Parmi les différents sens de l'imparfait que Labeau (2002) a répertoriés (cf. § 6.2, p. 193), son sens de base (principalement en vertu du critère de l'interprétation spontanée) est celui de simultanéité. L'imparfait a beaucoup fait parler de lui⁶⁹, mais en fin de compte ce qu'on en dit revient le plus souvent à constater qu'il exprime une sorte de «présent dans le passé», c'est-à-dire qu'il situe les faits par rapport à un point de référence qui n'est pas absolu et qui se trouve avant le moment d'énonciation⁷⁰. Le contexte doit donc permettre d'identifier par rapport à quel instant le fait dénoté par un verbe à l'imparfait est situé. L'imparfait pose les faits comme simultanés à ce point de référence passé (que nous notons T_1). Par exemple, pour la phrase *En 1998, elle travaillait à la banque* (ici, $T_1 =$ 'en 1998'):

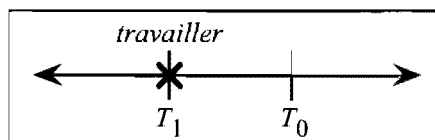


Figure 26 — Caractérisation sémantique de l'imparfait

Nous représentons 'travailler' par une croix dans cette figure, mais cela ne signifie pas que nous considérons le fait comme ponctuel. Dans ce cas-ci, la durée du fait dénoté par *travaillait* par rapport à l'an 1998 n'est pas pertinente: la personne en question peut avoir travaillé une partie de l'année à la banque, où encore y avoir travaillé pendant plusieurs années, y compris en 1998.

69. Voir notamment Le Goffic (1986a), Labeau & Larrivée (2005) ou Bres (2005).

70. Rappelons que nous avons rejeté plus haut (§ 6.1.2, p. 181) l'hypothèse d'un imparfait marquant l'aspect imperfectif.

Le sens aspectuel d'imperfectif qu'on peut percevoir dans cet exemple n'est pas, à notre avis, inhérent à l'imparfait. Il découle simplement de la perspective particulière qu'impose le point de référence, puisque le fait dénoté par le verbe est pour ainsi dire montré de l'intérieur.

6.3.5 *Le plus-que-parfait*

Le plus-que-parfait [*Tout à coup, elle avait pris son chapeau*] est construit avec AVOIR_{ant} et le suffixe –AI– de l'imparfait (cf. § 6.4.3, p. 206). Il est semblable à l'imparfait en ce qu'il ne situe pas les faits par rapport à un repère absolu. Il pose les faits comme étant antérieurs à un point de repère qui se trouve dans le passé (ici, T_1 n'est pas spécifié dans la phrase):

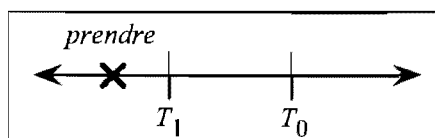


Figure 27 — Caractérisation sémantique du plus-que-parfait

Le plus-que-parfait n'a pas d'autre acception quand il est construit avec AVOIR_{ant}. Il ne faut pas le confondre avec le plus-que-parfait à valeur d'accompli, construit avec AVOIR_{acc}.

6.3.6 *Le conditionnel*

Enfin, le conditionnel, dans son acception strictement temporelle [*Ils ont dit qu'ils viendraient*], est l'image miroir du plus-que-parfait puisqu'il situe un fait comme postérieur à un repère passé. On remarquera que la situation temporelle du référent verbal par rapport au moment d'énonciation n'est pas pertinente. Les schémas de la Figure 28 ci-dessous représentent deux interprétations possibles de la même phrase (ici, $T_1 =$ 'dire').

Ce sens temporel peut se trouver transposé dans le temps, d'une façon similaire aux présent, imparfait et futur historiques déjà mentionnés: *Margarete entra à Ravensbruck le 2 août 1940. Elle n'en sortirait pas avant avril 1945* (Korzen & Nølke 1990: p. 277)⁷¹.

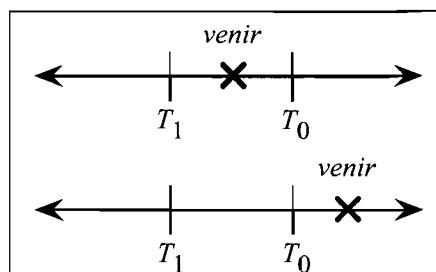


Figure 28 — Caractérisation sémantique du conditionnel

Le conditionnel est très souvent utilisé pour exprimer des sens modaux. Haillet (2002), notamment, dresse la liste de ces acceptions et les regroupe en deux classes de significations non-temporelles: le conditionnel d'hypothèse et le conditionnel «d'altérité énonciative». Nous ne visons pas la description exhaustive des acceptions modales que peut avoir le conditionnel, puisqu'elles ne sont pas directement pertinentes pour l'étude du temps grammatical. Nous nous contenterons de discuter trois de ces acceptions.

D'abord, le conditionnel peut exprimer une hypothèse liée à une condition non réalisée [*Si j'étais riche, je partirais en voyage*]. Ensuite, il peut exprimer une atténuation du propos [*Je reprendrais bien un peu d'anguille*]. Enfin, il exprime aussi la réserve du locuteur quand il rapporte des faits qu'on lui a relatés [*Le suspect serait entré par la fenêtre, selon les policiers*]. On trouve généralement de nombreux usages de cette acception dans les textes journalistiques.

Le conditionnel de réserve semble un bien meilleur candidat comme sens de base que le sens temporel que nous avons décrit plus haut (*cf.* § 3.4, p. 142). En effet, il est plus polyvalent syntaxiquement (il peut être le verbe principal de la phrase comme il peut dépendre d'un autre verbe, contrairement au conditionnel à valeur temporelle, qui ne peut apparaître que dans une subordonnée). Il répond aussi au critère de l'interprétation spontanée. Cependant, nous verrons que le conditionnel n'est pas un signe simple. Il est représentable, dans sa forme et dans son sens, par la composition de deux grammèmes: un qui indique la postériorité, et l'autre qui indique le décalage dans le passé du repère temporel. Ainsi, le sens tem-

71. Cité par Haillet (2002 : p. 24).

porel du conditionnel est parfaitement compositionnel, alors que ses sens modaux ne le sont pas.

Nous anticipons ici sur notre description des grammèmes temporels. Alors sans plus tarder, voyons quel système révèlent les caractérisations sémantiques sommaires que nous venons de faire des formes à valeur temporelle du français.

6.4 Le temps, c'est relatif

6.4.1 *Temps absolus, relatifs et absolus-relatifs*

De ce qui vient d'être dit en § 6.1, § 6.2 et § 6.3, on peut dégager une généralisation importante. Les formes verbales à valeur temporelle du français peuvent être divisées en deux classes: celles qui situent les faits directement par rapport au moment d'énonciation (le présent, le passé composé à valeur temporelle, le passé simple et le futur simple) et celles qui situent les faits par rapport à un repère temporel qui se trouve avant le moment d'énonciation (l'imparfait, le plus-que-parfait à valeur temporelle et le conditionnel). Il n'y a rien de nouveau ici, nous avons vu au Chapitre V que la dichotomie temps absolus ~ temps relatifs en français date au moins du XVIII^e siècle (*cf.* Ch. V, § 1, p. 72). Plus près de nous, cette opposition a aussi été théorisée en linguistique générale, par exemple par Comrie (1985) ou encore Mel'čuk (1994). Une hypothèse envisageable serait alors de stipuler deux catégories flexionnelles mutuellement exclusives, reprises au Tableau XXII ci-dessous:

- La catégorie de temps absolu, qui contiendrait les grammèmes correspondant au présent, au passé composé, au passé simple et au futur simple, qui situent les faits par rapport au moment d'énonciation.
- La catégorie de temps relatif, qui contiendrait les grammèmes correspondant à l'imparfait, au plus-que-parfait, au passé antérieur et au conditionnel, qui situent les faits par rapport à un autre fait dans le passé.

Toutefois, nous avons vu au Chapitre V (*cf.* Ch. V, § 1, p. 72) que cette hypothèse se heurte à un problème théorique si on tente de la traduire directement en

Temps absolus		Temps relatifs	
passé composé	<i>a aimé</i>	plus-que-parfait	<i>avait aimé</i>
passé simple	<i>aima</i>	imparfait	<i>aimait</i>
présent	<i>aime</i>	conditionnel	<i>aimerait</i>
futur	<i>aimera</i>		

Tableau XXII — Un système temporel classique à catégories concurrentes

termes de catégories flexionnelles et de grammèmes au sens où nous l'entendons (cf. Chapitre IV). Il faut forcément considérer ces catégories de temps relatif et de temps absolu comme mutuellement exclusives, puisqu'un fait ne peut pas être situé à la fois par rapport au moment d'énonciation et par rapport à un autre fait. Le locuteur doit donc choisir un grammème dans une seule des deux catégories stipulées, puisqu'un verbe ne peut pas recevoir un grammème des deux catégories à la fois. Il faut alors justifier que les verbes à l'indicatif soient parfois fléchis selon une catégorie et parfois selon une autre. Il nous semble difficile de formuler une règle élégante qui assure cela.

En plus des temps absolus et relatifs, Comrie (1985) stipule l'existence de temps absolus-relatifs, qui situeraient les faits par rapport à un point de référence qui lui-même serait situé par rapport au moment d'énonciation. Il s'agit en quelque sorte d'une reformulation de la théorie de Reichenbach (1947). Toutefois, Comrie considère cette «catégorie flexionnelle»⁷² comme complémentaire aux deux autres catégories de temps, c'est-à-dire qu'au lieu des deux catégories concurrentes de temps absolus et relatifs, on a chez Comrie trois catégories parallèles. On se retrouve donc avec le même problème: comment se fait-il que le locuteur ait le choix, pour tous les verbes à l'indicatif, de les fléchir soit en temps absolu, soit en temps relatif, soit encore en temps absolu-relatif?

Puisque les temps absolus et relatifs peuvent être considérés comme des types particuliers de temps absolus-relatifs en français (le moment de référence est sim-

72. Comrie ne parle pas vraiment de catégories flexionnelles. Ses catégories ne sont que des regroupements de signes ayant certaines caractéristiques communes.

plement simultan e au moment d' enonciation dans le cas des temps absolus, ou encore simultan e   un autre fait dans le cas des temps relatifs), on peut faire l' conomie de ces deux cat egories pour ne garder que celle de temps absolu-relatif. On aurait donc une seule cat egorie pouvant comporter en principe jusqu'  neuf gramm emes exprimant les sens 'simultan e   un rep ere pr esent (= simultan e au moment d' enonciation)', 'simultan e   un rep ere pass e', 'simultan e   un rep ere futur', 'ant erieur   un rep ere pr esent', 'ant erieur   un rep ere pass e', etc. Le syst eme verbal fran ais n'a toutefois pas de forme servant   situer explicitement un fait par rapport   un rep ere futur⁷³.

Ces sens grammaticaux th eoriques de temps absolu-relatif sont tous repr esentables en terme de deux s eries de trois sens: 'simultan e' ~ 'ant erieur' ~ 'post erieur' d'une part et 'par rapport   un rep ere pr esent ~ pass e ~ futur' d'autre part. Ce d ecoupage correspond aux sens que nous avons identifi es plus haut (§ 6.3, p. 195), que nous repr esentons sous forme de tableau ci-dessous (rappelons que T_0 repr esente le moment d' enonciation et que T_1 repr esente le rep ere temporel par rapport auquel est situ e T , c'est- -dire le moment ou le fait d enot e par le verbe se produit)⁷⁴:

	$T_1 = T_0$ -�-	$T_1 < T_0$ -ai-
$T < T_1$ AVOIR _{ant}	<i>mangea / a mang�e</i>	<i>avait mang�e</i>
$T \approx T_1$ -�-	<i>mange</i>	<i>mangeait</i>
$T > T_1$ -r-	<i>mangera</i>	<i>mangerait</i>

Tableau XXIII — Un d ecoupage des signes temporels du fran ais

73. Le futur ant erieur peut sembler un bon candidat [*Il aura mang e quand j'arriverai*], mais il s'agit d'un accompli. Ce qu'il d enote est la phase qui suit un fait, et c'est cette phase qui est situ e par rapport au moment d' enonciation.

74. Nos T et T_0 correspondent respectivement aux E et S de Reichenbach (*cf.* § 6.1.4, p. 187). Par contre, nous avons  mis des doutes sur la validit e de R (*cf.* p. 192), auquel ne correspond pas forc ement notre T_1 . Nous pr ef erons donc  viter la notation $E/R/S$.

On remarque que ce découpage sémantique va de paire avec un découpage formel (sauf pour le passé simple — nous reviendrons plus loin sur cette forme, Ch. VIII, § 3.4.1, p. 334). Nous avons donc affaire à cinq signes :

- $\langle T < T_1 \rangle \Leftrightarrow \text{AVOIR}_{\text{ant}}$
- $\langle T = T_1 \rangle \Leftrightarrow -\emptyset-$
- $\langle T > T_1 \rangle \Leftrightarrow -r-$
- $\langle T_1 = T_0 \rangle \Leftrightarrow -\emptyset-$
- $\langle T_1 < T_0 \rangle \Leftrightarrow -ai-$

Ils n'appartiennent toutefois pas tous à la même catégorie flexionnelle. Les trois premiers, qui situent les faits par rapport à un point de repère, font partie de la catégorie de temps à proprement parler. Les deux autres, qui situent le point de repère par rapport au moment d'énonciation, font partie de ce que nous appellerons la catégorie de « décalage ».

6.4.2 La catégorie flexionnelle de temps

La catégorie flexionnelle de **temps** contient trois grammèmes, qui sont des prédicats à deux arguments :

- Le grammème **antérieur** signifie 'X a lieu avant Y'. Il s'exprime généralement de façon analytique par la construction «AVOIR_{ant} + V-é». Dans le cas du passé simple, le grammème **antérieur** n'a pas de signifiant propre ; il est toujours exprimé dans le même morphe que la personne et le nombre.
- Le grammème **simultané** signifie 'X a lieu en même temps que Y'. Il s'exprime par un signe zéro, à savoir l'absence d'auxiliaire et de suffixe de temps.
- Le grammème **postérieur** signifie 'X a lieu après Y'. Il s'exprime de façon synthétique par le morphème -R- (le suffixe du futur simple).

N.B. : Idéalement, nous voudrions appeler les grammèmes de temps «passé», «présent» et «futur», au lieu de «antérieur», «simultané» et «postérieur». En effet, ces grammèmes correspondent à peu près à ceux qu'on trouve dans beaucoup d'autres langues, et l'usage de cette terminologie est très largement consensuel. Cependant, puisque nous continuons d'utiliser les termes traditionnels pour nommer les formes verbales («présent», «imparfait», «futur

simple», etc.), le risque de confusion est trop grand entre la forme du présent et le grammème **présent**, ou encore entre la forme du futur simple et le grammème **futur**. Des phrases comme «l'imparfait exprime le grammème **présent**» ou «le futur simple et le conditionnel expriment tous les deux le grammème **futur**» ne peuvent que donner des maux de tête au lecteur. Nous adoptons donc la terminologie «antérieur ~ simultané ~ postérieur».

Le temps est obligatoire à l'indicatif et ne peut pas s'utiliser sans un grammème de décalage. Il n'y a donc aucune forme verbale qui exprime uniquement les grammèmes **antérieur**, **simultané** ou **postérieur**; toutes les formes de l'indicatif expriment à la fois un grammème de temps et un grammème de décalage. Étant donné que le français oblige le locuteur à indiquer explicitement son repère temporel, il faut considérer les grammèmes de temps comme des prédicats sémantiques à deux arguments (le premier argument étant le fait à situer temporellement, et le second étant le repère par rapport auquel ce fait est situé). Dans les langues où les faits sont toujours situés par rapport au moment d'énonciation, on peut considérer les grammèmes de temps comme monoactanciels; en fait, leur «second actant» ne varie pas (c'est toujours 'maintenant') et il se trouve intégré à leur définition.

6.4.3 *La catégorie flexionnelle de décalage*

On sait depuis longtemps qu'en français, la série des temps absolus s'oppose à celle des temps relatifs notamment par le suffixe –AI– (celui de l'imparfait). Par exemple, Damourette & Pichon (1911–1950, tome V, p. 177) remarquaient que l'imparfait sert à exprimer «une autre actualité que celle où se trouve le locuteur au moment de la parole». Kahn (1954: p. 87) notait que «les variétés et les sous-variétés de l'imparfait correspondent jusqu'à un certain degré à celles du présent» et que «l'imparfait est par rapport à un moment-repère dans le passé ce que le présent est par rapport au moment de la parole». On l'a traditionnellement surtout remarqué pour l'opposition présent ~ imparfait, mais l'observation vaut aussi pour les oppositions futur ~ conditionnel ou passé composé ~ plus-que-parfait, comme le notait Vet (1980):

Il y a deux systèmes complets: le premier comportant le PC [= passé composé], le PR[ésent] et le FUT[ur], le second comportant le PQP [= plus-que-parfait], l'IMP[arfait] et le FUTP [= conditionnel]. Nous croyons que, du

point de vue sémantique, ces deux systèmes permettent de référer à deux mondes différents, le premier actuel, le second non actuel.

Vet (1980: p. 31)

Pourtant, très peu de linguistes ont essayé de décrire ces faits en termes d'une opposition de grammèmes au sein d'une catégorie flexionnelle. Martinet (1979) et Touratier (1996) l'ont fait:

Le monème décalé [...] indique que le procès exprimé dans le syntagme verbal est localisé selon l'axe du temps par rapport à un point de référence reporté — décalé — dans le temps révolu au moment de l'acte de parole.

Martinet (1979: p. 110)

Cependant, nous avons déjà vu que les catégories flexionnelles des fonctionnalistes ne répondent pas à nos critères puisqu'elles ne sont pas obligatoires (cf. Ch. V, § 3, p. 84). Nous proposons donc une catégorie qui respecte ces critères.

Ce que nous appelons la catégorie de **décalage** est la catégorie flexionnelle dont les grammèmes situent un repère temporel par rapport au moment d'énonciation. Ce repère est le point de vue duquel les faits dénotés par les verbes sont «montrés» par le locuteur. Cette catégorie flexionnelle peut compter en principe trois grammèmes (le repère pouvant être simultané, antérieur ou postérieur au moment d'énonciation). Nous verrons qu'il n'y en a que deux en français.

De façon provisoire, nous dirons que le grammème **non-décalé** signifie que le repère temporel est simultané au moment d'énonciation. Les grammèmes **antérieur**, **simultané** et **postérieur** situent alors les faits directement par rapport au moment présent. Le grammème **non-décalé** s'exprime par un suffixe zéro. On le retrouve dans les formes du présent, du passé composé à valeur d'antériorité, du passé simple et du futur simple. Ces formes ont justement en commun de poser les faits comme antérieurs, simultanés ou postérieurs *au temps de la parole*.

Le grammème **décalé** signifie que le repère temporel est dans le passé. Les grammèmes **antérieur**, **simultané** et **postérieur** situent alors les faits par rapport à ce point de repère passé. Le grammème **décalé** s'exprime par le suffixe –AI– (le suffixe de l'imparfait). On le retrouve dans les formes de l'imparfait, du plus-que-

parfait à valeur d'antériorité et du conditionnel (quand il a une valeur temporelle, bien entendu). Ces formes ont justement en commun de situer les faits comme antérieurs, simultanés ou postérieurs à *un point de repère passé*.

Le français ne dispose d'aucune forme permettant de situer explicitement un fait par rapport à un repère futur⁷⁵. Alors, que se passe-t-il quand un francophone cherche à situer un fait par rapport à un repère futur? Nous avons vu lors de notre discussion du modèle de Reichenbach (§ 6.1.4, p. 187) que la case vide du «futur postérieur» ne correspond à aucun temps spécifique en anglais, puisque c'est le futur simple qui remplit cette fonction (nous avons donné l'exemple *He will say that he will eat it the next day*). La même remarque vaut pour le français, qui marque le «futur dans le futur» par un futur simple :

(140) *Il dira que ce sera encore plus délicieux le lendemain.*

Il n'est pas possible de savoir si ce futur simple exprime un «futur dans le futur» ou simplement un futur par rapport au moment d'énonciation. Le fait que *sera* apparaisse dans une subordonnée peut porter à croire que le point de repère est *dira*, mais rien ne permet de l'affirmer hors de tout doute.

Si un fait est simultané à un repère futur plutôt que de lui être postérieur, c'est alors le présent qui doit être utilisé (le futur simple se comprend dans ce contexte comme indiquant la postériorité au repère). Ici, il est clair que c'est bien la concomitance à un point de repère futur qui est exprimée, et non la postériorité au moment d'énonciation :

- (141) a. *Il dira que c'est délicieux.*
 b. # *Il dira que ce sera délicieux.*

À la lumière de ces faits, on peut supposer que si un fait est antérieur à un repère futur, les formes à utiliser ne seront pas différentes de celles qui expriment l'antériorité par rapport à un repère présent. Effectivement :

75. Selon El-Kassas (2005), de telles formes existent en arabe standard.

- (142) a. *Il dira que ça a été délicieux.*
 b. *Il dira que ce fut délicieux.*
 c. # *Il dira que ça aura été délicieux.*

Mais comment expliquer le contraste de grammaticalité entre l'énoncé (c) de cette série et le (b) de la série précédente? Nous croyons qu'il est dû au fait que le futur antérieur ne peut avoir qu'une valeur d'accompli, c'est-à-dire qu'il est forcément construit avec AVOIR_{acc}, et non AVOIR_{ant}. On peut le vérifier en utilisant les indices de perfectivité de Vetters (1996) déjà mentionnés (§ 6.1.2, p. 181). En effet, si le futur antérieur pouvait exprimer l'antériorité (comme son nom le laisse entendre), on s'attendrait à ce qu'il soit compatible avec les indices de perfectivité, comme le sont le passé composé et le plus-que-parfait. Or, ce n'est pas le cas :

- (143) a. *Il a couru jusqu'à 5 h.*
 b. *Quand il est rentré, il nous a dit qu'il avait couru jusqu'à 5 h.*
 c. # *Quand il rentrera, il nous dira qu'il aura couru jusqu'à 5 h.*
 d. *Quand il rentrera, il nous dira qu'il a couru jusqu'à 5 h.*
- (144) a. *Il a couru de 5 h à 7 h.*
 b. *Quand il est rentré, il nous a dit qu'il avait couru de 5 h à 7 h.*
 c. # *Quand il rentrera, il nous dira qu'il aura couru de 5 h à 7 h.*
 d. *Quand il rentrera, il nous dira qu'il a couru de 5 h à 7 h.*
- (145) a. *Il a écrit cet article en 2 h.*
 b. *Quand il est rentré, il nous a dit qu'il avait écrit cet article en 2 h.*
 c. # *Quand il rentrera, il nous dira qu'il aura écrit cet article en 2 h.*
 d. *Quand il rentrera, il nous dira qu'il a écrit cet article en 2 h.*
- (146) a. *Il a sonné trois fois avant d'entrer.*
 b. *Quand il est rentré, il nous a dit qu'il avait sonné trois fois.*
 c. # *Quand il rentrera, il nous dira qu'il aura sonné trois fois.*
 d. *Quand il rentrera, il nous dira qu'il a sonné trois fois.*
- (147) a. *Tout à coup, il a viré de bord.*
 b. *Quand il est rentré, il nous a dit qu'il avait tout à coup viré de bord.*
 c. # *Quand il rentrera, il nous dira qu'il aura tout à coup viré de bord.*
 d. *Quand il rentrera, il nous dira qu'il a tout à coup viré de bord.*

On observe le même phénomène avec le «conditionnel passé» (nous excluons ici ses valeurs modales [*D'après Marie, il aurait couru jusqu'à 5 h*]):

- (148) a. *Nous savions qu'il nous dirait qu'il a couru jusqu'à 5 h.*
 b. # *Nous savions qu'il nous dirait qu'il aurait couru jusqu'à 5 h.*
 c. # *Nous savions qu'il nous dirait qu'il avait couru jusqu'à 5 h.*

Ces données nous permettent de tirer deux conclusions. Premièrement, le futur antérieur et le conditionnel passé ne peuvent avoir qu'une valeur d'accompli, c'est-à-dire qu'il n'existe pas en français d'«antérieur du futur» ou d'«antérieur du 'futur dans le passé'». C'est ce qui explique d'ailleurs que le futur surcomposé et le conditionnel surcomposé soient peu acceptés et ne puissent avoir qu'une interprétation modale (*cf.* § 4.1.3, p. 150 et § 4.1.4, p. 152), contrairement au passé surcomposé et au plus-que-parfait surcomposé qui expriment l'antérieur d'un accompli.

Deuxièmement, il faut en conclure que le grammème **non-décalé** peut également indiquer un repère temporel futur, et non seulement un repère présent. Il faut donc modifier notre définition pour dire que le grammème **non-décalé** signifie que le repère temporel n'est pas dans le passé (c'est-à-dire qu'il est soit simultané au moment d'énonciation, soit dans le futur).

La catégorie de décalage est obligatoire à l'indicatif et ne se combine pas aux autres modes. Elle est marquée sur la tête syntaxique de l'amas verbal. Elle ne s'exprime jamais seule, puisque la catégorie de temps est également obligatoire à l'indicatif. Il n'y a donc aucune forme verbale qui exprime simplement le grammème **non-décalé** ou **décalé**. En conséquence, l'imparfait n'est pas un grammème, mais est formé de la combinaison de deux grammèmes: **simultané** et **décalé**. Ainsi, nous avons pour le français un système temporel à deux dimensions, comme le montre le Tableau XXIV ci-dessous.

Ce modèle diffère des analyses en termes de «temps absolus» vs «temps relatifs» ou d'autres systèmes à paradigmes concurrents comme ceux de Imbs (1960), Weinrich (1973) ou Vet (1980). Le défaut principal de ces modèles est

		Décalage	
		non-décalé	décalé
Temps	antérieur	<i>a aimé / aima</i>	<i>avait aimé</i>
	simultané	<i>aime</i>	<i>aimait</i>
	postérieur	<i>aimera</i>	<i>aimerait</i>

Tableau XXIV — Notre système temporel à deux dimensions

qu'ils ne mettent pas en valeur les signes de la langue et ne permettent donc pas vraiment de bien expliquer la parenté sémantique frappante qu'il y a, par exemple, entre le présent et l'imparfait (*cf.* § 6.2, p. 193). Dans notre modèle, cette parenté s'explique tout naturellement par le fait que les deux portent le même grammème **simultané**. Les sens qui sont communs à ces deux formes sont en fait ceux de ce grammème (nous y reviendrons un peu plus loin, § 6.4.4, p. 213).

Notre système grammatical à deux catégories temporelles complémentaires permet également de rendre compte de l'opposition passé simple ~ imparfait (*cf.* § 6.1, p. 178). Il est vrai que les faits dénotés par les verbes au passé simple ou au passé composé sont perçus, sinon comme ponctuels, au moins comme complets, avec un début et une fin, alors que les faits dénotés par les verbes à l'imparfait sont plutôt perçus comme n'ayant pas de contours. Toutefois, nous croyons que cette impression est une illusion due à la perspective imposée par le sens de chacune de ces formes. Bien que toutes ces formes situent les faits dans le passé, elle ne le font pas de la même manière. Lorsqu'un fait est situé dans le passé par rapport à un repère présent, comme pour le passé composé et le passé simple, on montre pour ainsi dire le fait «de l'extérieur». C'est pour cette raison que ses bornes sont directement perçues. Au contraire, dans le cas de l'imparfait, le fait est présenté comme simultané à un point de référence passé. Le fait est donc vu «de l'intérieur». Ses bornes ne sont plus «visibles» et du coup le fait ne semble plus avoir de contours. C'est en quelque sorte comme regarder l'océan à partir d'un bateau au large: il revêt l'aspect d'une masse informe, dont on ne peut même pas dire s'il a un début et une fin. Au contraire, lorsqu'on l'observe de la côte, on le perçoit comme un objet aux contours bien définis. Ainsi, l'hypothèse que nous avons présentée per-

met d'expliquer de façon élégante la différence entre le passé composé et le passé simple d'une part (qui occupent la même case dans le Tableau XXIV) et l'imparfait d'autre part, tout en faisant l'économie d'une catégorie aspectuelle qui ne s'appliquerait qu'aux formes du passé. Le fait que le passé simple ne soit pas compositionnel dans sa forme n'est pas particulièrement gênant puisque son emploi est relativement restreint par rapport à son concurrent, le passé composé (probablement en partie, justement, à cause de son caractère non compositionnel).

Notre modèle explique également de façon toute naturelle le phénomène de la concordance des temps. Par exemple, considérons les phrases suivantes :

- (149) a. *Il pense qu'elle viendra.*
 b. *Il pensait qu'elle viendrait.*
 c. *Il pensera qu'elle viendra.*
- (150) a. *Il dit qu'elle est laide.*
 b. *Il a dit qu'elle était laide.*
 c. *Il dira qu'elle est laide.*

Le verbe subordonné situe un fait par rapport au moment où a lieu le fait dénoté par le verbe principal. Il porte alors le grammème de temps qui exprime la relation en question (**postérieur** pour les trois premières phrases, **simultané** pour les trois autres). Pour ce qui est du grammème de décalage, son choix dépend du moment où se déroule le fait dénoté par le verbe principal. Si ce dernier est dans le passé, comme dans les phrases (b) des deux séries, alors le verbe subordonné portera le grammème **décalé** puisqu'il est situé par rapport à un repère qui est dans le passé; autrement, il prendra le grammème **non-décalé**.

Il faut noter que quand le verbe de la subordonnée dénote un état qui transcende les époques (comme c'est le cas dans la deuxième série ci-dessus), alors le locuteur a le choix de situer cet état par rapport au référent du verbe principal ou par rapport au moment d'énonciation. Dans le cas où le verbe principal est au présent ou au futur, cela ne fait aucune différence visible, puisque c'est le même grammème **non-décalé** qui doit être utilisé, peu importe si le point de référence est actuel ou

ultérieur. Par contre, dans le cas d'un verbe principal au passé, on peut observer le phénomène :

- (151) a. *Il a dit qu'elle était laide.*
 b. *Il a dit qu'elle est laide.*

Il s'agit d'un choix du locuteur, selon qu'il souhaite prendre un point de vue décalé ou non, en fonction de ses buts communicatifs.

Nous illustrons le phénomène de la concordance des temps d'un point de vue formel Chapitre VIII (Ch. VIII, § 3.4, p. 334).

6.4.4 *Les acceptions secondaires des grammèmes de temps et de décalage*

Nous n'avons considéré jusqu'ici que les acceptions de base des grammèmes de temps et de décalage. Nous avons vu cependant que certaines formes verbales de l'indicatif peuvent avoir plus d'un sens (*cf.* § 6.2, p. 193 et § 6.3, p. 195), à savoir :

- le présent et l'imparfait de validité permanente, d'habitude, de passé récent, de futur inéluctable, historiques ou de condition ;
- le présent injonctif ;
- l'imparfait d'atténuation ou d'irréel ;
- le futur de supposition, historique ou d'atténuation ;
- le conditionnel d'hypothèse irréaliste, historique, d'atténuation ou de réserve.

Toutes ces formes portent un grammème de temps et un de décalage. Il faudra déterminer pour chacune de ces acceptions si elle est due à un seul de ces grammèmes ou à la combinaison des deux (en d'autres termes, il s'agit de savoir si nous avons affaire à un signe compositionnel ou à un phrasème grammémique). Nous allons commencer par les sens des formes du présent et de l'imparfait.

Les formes qui portent le grammème simultané

L'acception de base du grammème **simultané**, quand elle se combine à l'acception de base de **non-décalé**, donne ce que nous avons appelé plus haut le présent d'actualité [*Il pleut*]. Combinée à l'acception de base de **décalé**, cela donne l'imparfait d'actualité [*Il pleuvait*].

Nous avons vu plus haut (§ 6.2, p. 193) le parallèle frappant qui existe entre la plupart des autres sens du présent et de l'imparfait, qui se distinguent uniquement par le point de référence, c'est-à-dire par l'acceptation de base des grammèmes de décalage. La variété des sens parallèles de ces deux formes doit donc être due uniquement à diverses acceptations du grammème **simultané**. Voyons ce qu'il en est.

Le présent et l'imparfait de validité permanente [*Le soleil se couche <couchait> à l'Ouest*] ne se distinguent pas vraiment du présent et de l'imparfait d'actualité. Les énoncés de ce type expriment tout simplement des propositions qui sont tenues pour vraies au moment de référence spécifié par le grammème de décalage.

Le présent et l'imparfait d'habitude [*Il mange <mangeait> quatre fois par jour*] expriment un sens qui n'est pas propre au grammème **simultané**. En effet, on trouve le même sens, par exemple, à l'infinitif quand il est placé dans un contexte approprié [*Manger quatre fois par jour est bon pour la santé*]. Nous y voyons plutôt un signe n'ayant rien à voir avec le temps qui exprime le sens 'habituel'.

Le présent et l'imparfait de passé récent [*Je rentre <rentrais> tout juste de Berlin*] relève à notre avis d'un procédé narratif par lequel le locuteur fait semblant soit que le moment d'énonciation est déplacé dans le passé, soit que le fait dénoté par le verbe est en train de se produire. Ce qu'il exprime, c'est bien un présent ou un imparfait d'actualité, comme si le fait était vraiment actuel, mais le contexte permet de comprendre qu'il s'agit d'un passé. Nous croyons que ce phénomène ne doit pas être décrit par une grammaire.

Le présent et l'imparfait de futur inéluctable [*Nous partons <partions> à 5 h*] se traitent de la même façon que le présent et l'imparfait de passé récent. Il ne s'agit pas d'une véritable acceptation du grammème **simultané**, mais d'un procédé narratif.

Le présent et l'imparfait historiques [*Le 2 octobre 1535, Cartier arrive <arrivait> à Hochelaga*], encore une fois, ne relève pas de la grammaire au sens strict. Il s'agit d'un procédé narratif, une sorte de «mise en scène» du locuteur, qui fait

comme si le moment d'énonciation même se trouvait déplacé dans le temps. Le présent et l'imparfait historiques n'existent pas forcément dans toutes les langues. Par exemple, en russe, la traduction littérale de *Le 2 octobre 1535, Cartier arrivait à Hochelaga* (*2-go oktjabrja 1535-go goda Kart'e pribyval v Oshlagu) est absolument incorrecte. Les procédés narratifs en jeu ici varient d'une langue à l'autre et ils doivent être pris en charge par des règles spécifiques à une langue donnée, mais ces règles ne font pas partie du modèle que nous sommes en train de construire.

Le présent et l'imparfait de condition [*Elle dit <disait> que s'il pleut <pleuvait>, on ira <irait> au cinéma*] ne mettent pas en jeu une acception secondaire du grammème **simultané**. Les grammairiens remarquent que «après *si* conditionnel, on emploie obligatoirement le présent pour un fait futur» (Grevisse 1993: § 850). On en a souvent conclu que le présent peut avoir un sens de condition future. Mais ce n'est pas le cas. D'abord, le sens de condition provient de SI, qui ne sert à rien d'autre qu'à exprimer celui-ci. Ensuite, la proposition à vérifier n'est pas forcément dans le futur. Il peut très bien s'agir d'une proposition actuelle dont le locuteur ne connaît pas la valeur de vérité [*Si tu es riche, tu peux te le payer*]. Nous considérons que le grammème **simultané** est imposé par la combinatoire de SI, qui exige que son dépendant porte ce grammème. Le grammème **simultané** garde ici son sens premier de 'simultané'. Le point de référence peut cependant être situé n'importe où dans le temps, ce qui se reflète dans les grammèmes **non-décalé** [*Maintenant, si tu as de l'argent, tu peux te le payer* ou *Demain, si tu as de l'argent, tu pourras te le payer*]⁷⁶ et **décalé** [*Au XX^e siècle, si tu avais de l'argent, tu pouvais encore te payer une île*].

Voilà pour les acceptions parallèles du présent et de l'imparfait. Nous avons vu (§ 6.2, p. 193) qu'il existe aussi un sens du présent qui n'a pas d'équivalent à l'imparfait, et deux acceptions de l'imparfait qui ne trouvent pas d'écho au présent.

Le présent d'injonction [*On se calme!*] exprime un ordre. Il nous semble cependant que dans cette forme, les grammèmes **simultané** et **non-décalé** retiennent

76. Rappelons que non-décalé peut indiquer un point de repère futur (cf. § 6.4.3, p. 206).

leur sens premier. Il n'y pas d'équivalent à l'imparfait, mais cela est probablement dû à une incompatibilité sémantique. En effet, il ne fait aucun sens de faire une injonction décalée dans le passé, puisque cette époque est révolue et ne peut donc plus être influencée par un acte de parole. Le présent injonctif ne peut s'utiliser qu'en discours direct, et seulement dans la proposition principale [[#]*Il dit qu'on se calme!*]. C'est ce qui nous fait croire que le signifiant du signe en jeu ici n'est pas la forme du présent, mais la construction dans laquelle elle apparaît, ainsi qu'un prosodème (ce sens ne peut être exprimé qu'avec une intonation particulière). Ce phénomène linguistique ne relève donc pas des grammèmes de temps ou de décalage.

L'imparfait d'atténuation [*Je voulais vous demander quelque chose*]⁷⁷ n'a pas d'équivalent au présent. Il est très proche sémantiquement du conditionnel d'atténuation [*Je voudrais vous demander quelque chose*], dont nous parlerons plus loin. Ce qu'il exprime, c'est en fait un présent plus l'atténuation. Donc, le grammème **simultané** exprime ici son sens premier. C'est le grammème **décalé** qui exprime le sens d'atténuation. Nous avons donc une acception secondaire pour ce grammème. Son signifié ne se laisse pas bien décrire par une définition, puisqu'il relève plutôt de la structure rhétorique, où sont traités les phénomènes liés au registre et à la relation entre les participants de la situation de communication.

L'imparfait d'irréalité [*Si j'étais riche, je partirais en voyage autour du monde*] n'a pas non plus d'équivalent au présent. Nous avons déjà dit plus haut que nous croyons que *SI* impose le grammème **simultané** à son dépendant. Ce qui distingue *si tu es riche* de *si tu étais riche*, c'est seulement le grammème de décalage (les deux portent le grammème **simultané**, qui, dans les deux cas, exprime son sens premier). Le grammème **décalé** présente ici la proposition conditionnelle comme irréelle. Il s'agit là d'une autre acception de **décalé**, qui ne se laisse pas bien définir par une décomposition sémantique. Son signifié relève de la structure référentielle, qui lie les sémantèmes aux référents du monde réel vers lesquels ils pointent.

77. Nous considérons l'imparfait dit « forain » [*Qu'est-ce qu'il lui fallait, à la petite dame?*] et autres emplois hypocoristiques comme des cas d'imparfait d'atténuation.

En bref, malgré la variété des emplois du présent et de l'imparfait, le grammème **simultané** n'a qu'une seule acception. Idem pour **non-décalé**. Par contre, le grammème **décalé** en a trois: son sens de base (qui indique que le point de repère temporel est décalé dans le passé), un sens d'atténuation (qui relève de la structure rhétorique) et un sens d'irréalité (qui relève de la structure référentielle).

Les formes qui portent le grammème postérieur

Passons maintenant aux formes qui portent le grammème **postérieur**, soit le futur simple et le conditionnel. Ces deux formes, dans leur emploi de base, se distinguent par le décalage (**décalé** pour le conditionnel, **non-décalé** pour le futur simple). Outre leur emploi de base, ils peuvent exprimer des sens modaux, comme nous l'avons dit plus haut. Voyons de quoi il retourne.

Le futur et le conditionnel dits «historiques» [*Le 15 mai 1940, l'état-major néerlandais capitulera <capitulerait> sans condition*] doivent être traités de la même façon que le présent et l'imparfait historiques, c'est-à-dire comme une mise en scène où le locuteur déplace virtuellement le moment d'énonciation, et non comme une acception particulière du grammème **postérieur**.

Le futur simple et le conditionnel partagent la propriété de pouvoir se retrouver dans le second terme d'une condition [*Si tu as <avais> de l'argent, tu pourras <pourrais> y aller*]. Il y a cependant des différences entre les deux. Le conditionnel, dans cette position, ne peut s'utiliser avec le même sens que si la condition n'est pas réalisée⁷⁸ [*?#Si tu as de l'argent, tu pourrais y aller*]⁷⁹. Il ne peut pas commuter avec l'imparfait [*#Si tu avais de l'argent, tu pouvais y aller*]⁸⁰, alors que le futur simple commute dans cette position avec le présent [*Si tu as de l'argent, tu peux y aller*] avec une différence sémantique prévisible. L'usage du futur simple et du conditionnel dans les propositions conditionnelles n'est donc pas tout à fait parallèle. Le futur simple, dans ces constructions, exprime son sens premier. Le conditionnel exprime le grammème **décalé** à valeur d'irréalité déjà mentionné plus

78. Cf. le grammème **décalé** à valeur d'irréalité dans la condition [*Si tu avais...*].

79. Cette phrase peut être grammaticale, mais il s'agit alors du conditionnel d'atténuation.

80. Cette phrase est grammaticale si l'imparfait *avais* n'exprime pas l'irréalité

haut, ainsi que ce qui semble être une nouvelle acception du grammème **postérieur**; en fait, ce grammème de temps est plutôt imposé par la construction et il est vide de sens.

Le futur et le conditionnel d'atténuation [*Je ne vous cacherai pas que j'ai eu peur; Je reprendrais bien un peu d'anguille*] expriment un sens très semblable à l'imparfait d'atténuation, qui relève de la structure rhétorique. Ce sens est exprimé par le grammème **postérieur**; il s'agit d'une acception secondaire. Le futur et le conditionnel d'atténuation expriment tous les deux cette même grammaire. Le conditionnel porte en plus le même grammème **décalé** que dans l'imparfait d'atténuation. Ce conditionnel est donc compositionnel: il exprime à la fois un futur d'atténuation et un imparfait d'atténuation. Nous ne sommes pas en mesure d'expliquer la nuance qui peut exister entre ces acceptions de **postérieur** et de **décalé**. Nous sommes loin des sens temporels qui nous intéressent dans cette thèse.

Le futur simple peut également avoir un sens modal indiquant une supposition de la part du locuteur [*J'entends des pas. Ce sera sans doute Lucie*]. Ce sens, qu'on peut décrire par 'je suppose que X', est propre au futur simple et ne trouve pas son équivalent au conditionnel. C'est donc la combinaison **postérieur** ⊕ **non-décalé** qui exprime ce sens modal, et non seulement **postérieur**. Il s'agit d'un phrasème grammématique (cf. Ch. IV, § 7, p. 67).

Enfin, le conditionnel peut servir à exprimer la réserve du locuteur quand il rapporte des faits qu'on lui a relatés [*Le suspect serait armé et dangereux, d'après les policiers*]. Nous représentons le sens de cette acception par 'il paraît que X'. On en trouve généralement de nombreux usages dans les textes journalistiques. Là encore, ce sens n'est attribuable ni à **postérieur**, ni à **décalé**, mais à la combinaison des deux. C'est un autre phrasème grammématique.

Nous avons délibérément laissé de côté les signes de phase aspectuelle, afin de ne pas confondre des phénomènes linguistiques très proches. Les marqueurs de phases ne montrent pas de polysémie, mais il existe un phrasème grammématique

impliquant l'auxiliaire de la phase accomplie: le futur antérieur à valeur de bilan [*L'hiver 1528–1529 aura été particulièrement rigoureux*]. Le sens de «*X aura V-é*» est plus ou moins 'en rétrospective, X a V-é'. Il contient donc le sens de AVOIR_{acc}, mais pas celui de **postérieur**. Il s'agit d'un cas intéressant de phraséologie, qu'on pourrait appeler **collocation grammaticale**⁸¹.

6.5 Synthèse: les catégories flexionnelles de temps et de décalage

Il y a donc en français non pas une, mais deux catégories flexionnelles distinctes qui situent les faits dans le temps. Une première, celle de décalage, indique quel est le repère temporel par rapport auquel sont situés les faits. Une deuxième catégorie, celle du temps à proprement parler, situe les faits par rapport à ce point de repère.

Les grammèmes de la catégorie de décalage sont:

- **Non-décalé**, qui signifie que le point de repère est soit le moment d'énonciation, soit un moment dans le futur. Il s'exprime par un suffixe zéro.
- **Décalé**, qui indique que le repère temporel est dans le passé. Il s'exprime par le suffixe *-AI-* (dont les allomorphes sont *-ai-* [*faisait*] et *-i-* [*faisions*]). Il peut également avoir une valeur d'atténuation ou d'irréalité.

La catégorie flexionnelle de temps compte les grammèmes suivants:

- **Simultané**, qui signifie 'X a lieu en même temps que Y'. Il s'exprime par un suffixe zéro.
- **Antérieur**, qui signifie 'X a lieu avant Y'. Il s'exprime généralement de façon analytique par la construction «AVOIR_{ant} + V-é». Quand il est combiné au grammème **non-décalé**, dans un registre soutenu ou littéraire, il peut aussi s'exprimer par le suffixe du passé simple, qui exprime de façon cumulative le temps, la personne et le nombre.
- **Postérieur**, qui signifie 'X a lieu après Y'. Il s'exprime de façon synthétique par le morphème *-R-*. Il peut également avoir une valeur d'atténuation.

81. Une collocation est une expression phraséologisée dont le sens comprend dans une position centrale celui d'une seule de ses composantes. Nous y reviendrons au Chapitre VIII (§ 3.5, p. 341).

Aucune forme verbale ne porte seulement qu'un grammème de temps ou de décalage. Ces deux catégories flexionnelles doivent toujours être utilisées ensemble.

7 Le genre, le nombre et la personne

Tous les verbes finis doivent s'accorder en personne et en nombre avec leur sujet. Les verbes au participe passé doivent quant à eux s'accorder en genre et en nombre (soit avec le sujet, soit avec l'objet, selon le contexte — nous y reviendrons plus loin, Ch. VIII, § 2.4.5, p. 310). Notre traitement de ces catégories flexionnelles suit la tradition. La catégorie flexionnelle de genre ne contient que les grammèmes **masculin** et **féminin**. La catégorie de personne contient les grammèmes **1^{ère}**, **2^e** et **3^e**. Les grammèmes de nombre sont **singulier** et **pluriel**.

Les catégories flexionnelles de genre, de nombre et de personne sont des catégories d'accord, et en ce sens, leurs grammèmes ne sont pas porteurs de sens. Ces signes grammaticaux ont peu d'intérêt pour notre étude et nous nous contenterons de cette description sommaire. Nous en reparlerons plus loin lors de la formalisation de notre modèle (Ch. VIII, § 2.4.5, p. 310 et Ch. VIII, § 3.2.1, p. 319).

8 Synthèse: les catégories flexionnelles et leurs grammèmes

Les formes surcomposées nous ont permis de mettre en lumière le fait qu'il existe deux auxiliaires **AVOIR** en français: **AVOIR_{acc}** et **AVOIR_{ant}**. Le premier fait partie des marqueurs de phases, qui ne forment pas une catégorie flexionnelle à proprement parler. Les marqueurs de phases expriment des sens liés aux bornes temporelles des faits et aux phases qu'elles délimitent. L'auxiliaire **AVOIR_{acc}** exprime la phase qui suit l'accomplissement d'un fait. La locution prépositive **EN TRAIN** exprime la phase simultanée à la réalisation du fait. L'auxiliaire **ALLER** dénote la phase qui précède un fait. D'autres signes servent à dénoter les bornes temporelles des faits ou leur environnement immédiat, notamment **VENIR** et **SUR LE POINT**. Parmi tous ces marqueurs, ceux qui sont de type verbal (**ALLER**, **VENIR**

et AVOIR_{acc}) doivent, comme tous les autres verbes, porter des grammèmes de finitude, de mode, de temps, etc., ce qui permet de situer dans le temps les phases qu'ils dénotent.

La catégorie de finitude est obligatoire pour tous les verbes. Elle comprend les grammèmes **fini**, **infinitif**, **participe-ant** et **participe-é**.

Les verbes finis sont les seuls à devoir porter un grammème de mode. La catégorie de mode comprend les grammèmes **indicatif**, **subjonctif** et **impératif**. Nous considérons les deux premiers comme sémantiquement vides dans leur emploi de base, alors que l'impératif exprime un ordre, une requête, etc.

Seuls les verbes à l'indicatif permettent de situer les faits dans le temps sans avoir nécessairement recours à un marqueur de phase. Cela se fait *toujours* par le biais de deux catégories flexionnelles complémentaires. La catégorie de décalage indique quel est le point de repère par rapport auquel sont situés les faits (**non-décalé** si le point de repère est le moment de la parole ou un moment futur, **décalé** si le repère temporel est dans le passé). La catégorie de temps situe les faits par rapport à ce point de repère (**antérieur**, **simultané**, **postérieur**).

La Figure 29 ci-dessous montre de façon schématique les catégories flexionnelles verbales du français et la combinatoire de leurs grammèmes superficiels.

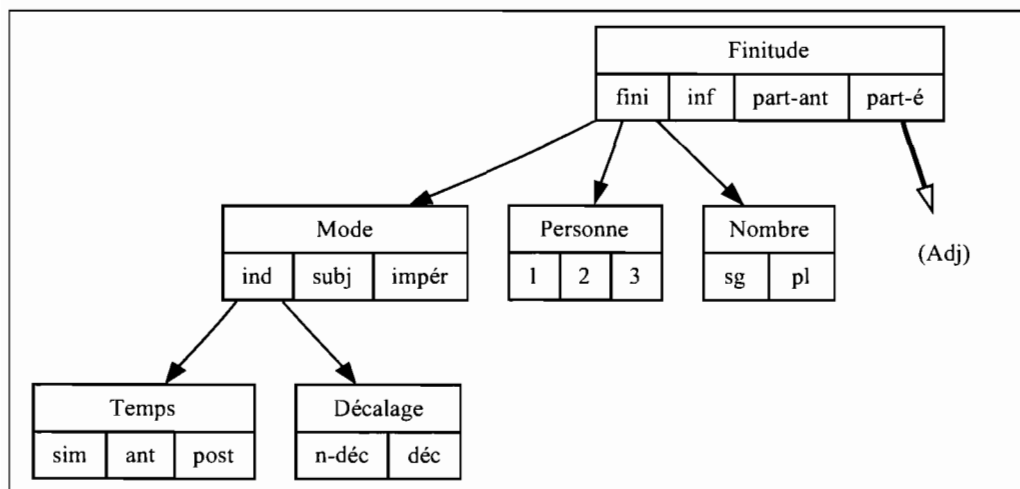


Figure 29 — La combinatoire des grammèmes superficiels verbaux en français

VI. Un modèle descriptif de la conjugaison française

Nous avons obtenu ce modèle en ne considérant que le sens de base des grammèmes. Autrement, nous n'aurions pas pu obtenir un système cohérent. Il est intéressant de noter que les acceptions secondaires des grammèmes et les phrasèmes grammémiques permettent d'exprimer des sens typiques de catégories flexionnelles qu'on trouve dans d'autres langues. Notamment, le conditionnel de réserve et le futur de supposition correspondent respectivement aux grammèmes **citatif** et **présomptif** de la catégorie d'évidentialité de Mel'čuk (1994: p. 167). On peut dire en quelque sorte que l'évidentialité est une «catégorie flexionnelle virtuelle» du français.

Nous allons maintenant voir comment formaliser notre description du système flexionnel verbal du français.

VII. La grammaire d'unification Sens-Texte

Dans les chapitres précédents, nous avons exposé notre méthodologie pour l'étude des grammèmes et l'avons appliquée au système verbal du français. Nous en avons dégagé un système flexionnel cohérent, mais nous ne l'avons décrit que de façon informelle. Nous allons maintenant voir comment ce modèle peut être formalisé. Notre but est de construire un dispositif formel qui mette en correspondance la représentation du sens d'un énoncé (avec les informations sémantiques relatives au temps) et la représentation de l'organisation syntaxique de cet énoncé, et ce, que ce soit dans le sens de la synthèse ou de l'analyse. Rappelons que notre travail se situe dans le cadre de la théorie Sens-Texte (TST), que nous avons présentée au Chapitre III. Cependant, le formalisme que nous utilisons est celui des grammaires d'unification Sens-Texte (GUST).

Le formalisme GUST a d'abord été proposé par Kahane (2001, 2002, 2003b) comme une reformulation de la TST en termes d'une grammaire basée sur l'unification de structures. Puis, Kahane (2004) a suggéré l'utilisation de polarités pour contrôler la saturation des structures. Construisant sur cette idée, Kahane & Lareau (2005a, 2005b) ont donné à GUST sa forme actuelle en précisant le formalisme et en élaborant le mécanisme d'articulation des modules de la grammaire. Dans ce chapitre, nous présenterons en détail GUST en reprenant notamment trois articles que nous avons publiés, soit Kahane & Lareau (2005a, 2005b) et Lareau (2007). Une partie significative de notre recherche sur le sujet a été faite en étroite collaboration avec Sylvain Kahane, si bien qu'il n'est pas toujours évident de savoir qui a fait quoi exactement. Cependant, étant donné la nature de cette thèse, nous tâcherons dans la mesure du possible de préciser notre rôle dans l'élaboration du formalisme.

1 L'architecture générale de GUST

L'architecture de GUST est semblable à celle de la Théorie Sens-Texte (*cf.* Chapitre III, p. 35). La principale différence entre les deux réside dans le fait que la TST présuppose une procédure implicite pour gérer l'interaction des règles. Il faut adjoindre à la grammaire un mécanisme qui vérifie quel fragment de structure a déjà été « consommé » par une règle, afin d'éviter d'une part que plusieurs règles s'appliquent au même fragment et génèrent des structures mal formées, et afin de vérifier d'autre part que toute la structure de départ a été « consommée » par les règles¹. GUST, quant à lui, se base sur l'unification de structures pour la combinaison des règles. Il s'apparente donc, en ce sens, aux grammaires d'arbres adjoints (TAG, de l'anglais *tree adjoining grammar*) (Joshi *et al.* 1975), aux grammaires syntagmatiques dirigées par les têtes (HPSG, de l'anglais *head-driven phrase structure grammar*) (Pollard & Sag 1994) et aux grammaires lexicales fonctionnelles (LFG, de l'anglais *lexical functional grammar*) (Bresnan 2001). Par ailleurs, la saturation des objets est explicitement prise en charge par les règles elles-mêmes grâce à un système de polarités (Kahane 2004, Kahane & Lareau 2005a), que nous présentons plus loin (§ 3, p. 229). Ces deux propriétés doivent en théorie rendre GUST plus facile à vérifier et à implémenter que la TST, puisqu'il n'y a pas de procédure cachée à prendre en charge.

En GUST, comme pour la TST, la correspondance entre un sens et un texte est décrite via un certain nombre de correspondances intermédiaires. Alors que la TST stipule sept niveaux de représentation (*cf.* Ch. III, § 2, p. 38), GUST n'en compte que quatre, puisqu'il ne distingue pas les niveaux profonds et de surface. Il s'agit des niveaux sémantique, syntaxique, morphotopologique et phonologique. Les structures utilisées diffèrent à chacun des niveaux de représentation : la structure sémantique est un graphe, la structure syntaxique est un arbre de dépendances non ordonné, la structure morphotopologique en est une de type syntagmatique où les

1. Bohnet & Wanner (2001) décrivent cette procédure. Nous avons aussi présenté un mécanisme similaire dans notre mémoire de maîtrise (Lareau 2002). Il faut noter que certaines formalisations de la TST disent explicitement ce qui est consommé par les règles (*cf.* notamment Kahane & Mel'čuk 1999 et Bohnet & Wanner 2001).

éléments sont ordonnés et la structure phonologique est une chaîne. Chaque niveau possède sa propre grammaire de bonne formation, qui assure la construction de structures valides. Puisque ces grammaires ne font que construire des représentations linguistiques sans les mettre en correspondance, il faut également des grammaires d'interface pour assurer cette mise en correspondance. On a donc quatre grammaires de bonne formation et trois grammaires d'interface, que nous regroupons toutes sous le terme générique de «module». Cette architecture très modulaire annonce une philosophie «atomiste», héritée de la TST, qui sous-tend cette théorie. En effet, GUST tend à éclater le problème de la correspondance entre les sens et les textes en sous-problèmes.

Il peut alors paraître paradoxal que GUST élimine les niveaux profonds de représentation utilisés dans la TST. En fait, Kahane (2002, 2003c) montre que le niveau syntaxique profond de la TST est équivalent à une structure de dérivation où apparaissent les signes profonds qui établissent la correspondance entre une représentation sémantique et une représentation syntaxique de surface données. Si on considère que chaque règle de l'interface sémantique-syntaxe de GUST décrit un signe profond, alors on peut dire que cette interface correspond au niveau syntaxique profond de la TST. On suppose que la même équivalence existe entre les autres modules d'interface de GUST et les autres niveaux profonds de la TST, bien que cela n'ait pas été discuté en profondeur dans la littérature. Dans cette thèse, nous ne nous intéressons pas aux niveaux de représentation morphologique et phonologique; nous laisserons donc cette question en suspens.

La TST a toujours beaucoup mis l'emphase sur la description du lexique dans un dictionnaire et a développé des outils spécifiquement adaptés à cette tâche (tableaux de régime, fonctions lexicales, etc.). En GUST, il n'y a pas de dictionnaire: tout est décrit dans la grammaire. Il y a donc forcément beaucoup de règles (en fait, l'écrasante majorité) qui font directement référence à un sémantème, un lexème ou un morphème particulier. En apparence, il s'agit d'une différence fondamentale entre la TST et GUST. Pourtant, il n'en est rien. En fait, le dictionnaire de la TST peut être conçu comme version «compactée» d'un ensem-

ble de règles ayant en commun une référence directe à un même lexème (ou aux sémantème et morphème qui lui correspondent); on a peu insisté sur cette vision des choses dans la littérature, mais elle n'est pas du tout incompatible avec la théorie. Par exemple, Kahane & Polguère (2001) ont démontré que les fonctions lexicales peuvent être définies formellement par des patrons de correspondance entre des structures sémantiques et syntaxiques profondes. Donc, il est possible de représenter les fonctions lexicales par des (patrons de) règles de correspondance. Il en va de même des patrons de régime et des autres informations contenues dans un article de dictionnaire (sauf, bien entendu, les informations métalinguistique, comme les exemples)². C'est du moins le postulat adopté en GUST, où tout le lexique est décrit par des règles. Il va de soi que cela augmente considérablement le nombre de règles à gérer dans la création d'une grammaire à large couverture, mais on peut très bien imaginer des outils informatiques qui faciliteraient cette tâche.

Nous verrons plus loin (§ 5, p. 236) quel est le rôle exact de chacun des modules de GUST et comment ils s'articulent pour former une grammaire complète. Mais d'abord, il convient de présenter en détail le formalisme utilisé.

2 La forme des structures de GUST

Chaque module de GUST est un ensemble non vide de règles. Contrairement aux règles de la TST, celles de GUST ne sont pas toutes des règles de correspondance. Elles sont toutes des **structures élémentaires** qui peuvent se combiner entre elles pour former des structures plus complexes qui représentent les énoncés linguistiques à divers niveaux de fonctionnement de la langue³. Par exemple, les deux structures suivantes peuvent constituer des règles en GUST (elles sont de types différents et appartiennent à des modules distincts de la grammaire). Il est normal que les détails du formalisme échappent au lecteur à ce stade-ci de l'exposé, mais on comprend intuitivement que la première règle représente le fait

2. Sur la transformation automatique d'articles de dictionnaire en règles dans le cadre de la TST, voir notamment Lareau (2002).

3. Formellement, rien ne distingue les règles d'un module des structures servant à représenter les énoncés. Les règles de GUST ne sont rien d'autre que des structures.

que ‘dormir’ est un prédicat à un argument et que la seconde indique que ce prédicat peut s’exprimer par le lexème DORMIR.

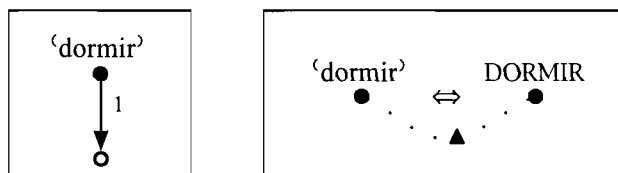


Figure 30 — Deux règles simples de GUST

En GUST, il n’y a aucune différence formelle entre une structure et une règle. Une même règle peut se combiner plusieurs fois avec d’autres, ou encore avec elle-même. C’est en combinant les règles, autrement dit en «collant» des bouts de structures ensemble, qu’on construit les représentations (sémantiques, syntaxiques, etc.) des énoncés, ainsi que les liens de correspondance entre elles.

Parmi les structures élémentaires qui forment un module, on peut en identifier une ou plusieurs comme **structures initiales**. Dans ce cas, lors de la combinaison des structures élémentaires du module, une des structures initiales doit être utilisée une et une seule fois. Nous verrons plus loin que les structures initiales sont notamment utilisées en syntaxe (cf. Figure 42, p. 246, et Figure 44, p. 249).

Toutes les structures sont conçues comme un ensemble d’**objets** et un ensemble de **fonctions** sur ces objets⁴. Les objets sont typés : nœud sémantique, arc syntaxique, objet grammémique, etc. Tout, dans une règle GUST, est un objet, y compris les liens de correspondance entre des nœuds de niveaux adjacents⁵. Les types d’objets sont décidés par le linguiste, selon le genre de grammaire qu’il souhaite construire⁶. Nous reviendrons plus loin sur les types d’objets que nous utilisons ; pour l’instant, il suffit de savoir qu’il existe différents types.

-
4. Il s’agit de fonctions du type $f(x) = y$ (voir la remarque à ce sujet, p. 229). Cette représentation des structures en termes d’objets et de fonctions sur les objets est notre contribution personnelle.
 5. Dans la formulation initiale de GUST (Kahane 2002), les liens de correspondance étaient implicites dans le formalisme.
 6. Kahane (2004) montre que le formalisme des grammaires d’unification polarisées permet de modéliser diverses théories linguistiques.

Les objets peuvent être associés à certaines fonctions, selon leur type. Par exemple, un nœud syntaxique pourra être associé à des fonctions qui retournent comme valeur un objet grammémique, alors qu'un arc ne peut pas être lié à de telles fonctions (nous verrons plus loin quelles sont les fonctions qui s'associent à tel ou tel type d'objet). Les objets ne sont pas intrinsèquement liés les uns aux autres ; ce sont plutôt les fonctions qui fournissent la structure proprement dite en liant les objets. Toutes les fonctions prennent un seul argument, toujours un objet d'une structure, et ne retournent qu'une seule valeur. Il existe trois types de fonctions, selon le type de valeur retournée :

- 1) Les **fonctions structurantes** associent un objet à un autre objet de la même structure. Elles servent, par exemple, à associer un arc à son nœud cible et à son nœud source, ou encore à rattacher les objets grammémiques aux nœuds syntaxiques. Ce sont ces fonctions qui font qu'une structure est bien une structure, et non un simple ensemble d'objets. Par exemple, les fonctions *source* et *cible* à la Figure 31 (p. 229) sont des fonctions structurantes qui lient un arc sémantique à deux nœuds sémantiques.
- 2) Les **fonctions d'étiquetage** associent à un objet une chaîne de caractères. Elles servent à étiqueter les objets de noms de lexèmes, de grammèmes, de relations syntaxiques, etc. Les étiquettes peuvent être spécifiques à une langue en particulier. Par exemple, la fonction *étiquette* à la Figure 31 (p. 229) est une fonction d'étiquetage. Elle lie un arc sémantique à un nom de relation sémantique, «1». Elle lie également un nœud sémantique à un sémantème, 'dormir'. La liste des relations sémantiques est universelle mais la liste des sémantèmes change d'une langue à l'autre.
- 3) Les **fonctions de polarisation** retournent comme valeur un élément d'un ensemble de polarités. Nous reviendrons à la section suivante sur la polarisation des objets et nous verrons un exemple de fonction de polarisation à la Figure 32 (p. 230).

N.B. : Nos «fonctions» ne sont pas tout à fait des fonctions au sens mathématique du terme⁷. En mathématiques, si une fonction f associe toujours à un même objet les mêmes valeurs que la fonction g , alors f et g désignent la même fonction. Dans le cas de GUST, deux fonctions qui associent les mêmes valeurs aux mêmes objets sont bien des fonctions distinctes si elles ne portent pas le même nom. Ainsi, contrairement aux fonctions en mathématiques, que l'on entend comme un ensemble de couples (*argument*, *valeur*), les fonctions de GUST sont plutôt des triplets (*nom de fonction*, *argument*, *valeur*). Elles sont en ce sens similaires aux fonctions des langages de programmation C ou Java par exemple, ou encore à des prédicats logiques binaires.

Ce formalisme permet par exemple d'encoder la structure sémantique qui représente le sens '[X] dormir' de la façon suivante. Nous avons trois objets, identifiés par les numéro 1 à 3. Chaque objet est typé. Aux objets s'ajoutent des fonctions. On a donc le couple $\langle \{ 1:\text{nœud}_{\text{sém}}, 2:\text{nœud}_{\text{sém}}, 3:\text{arc}_{\text{sém}} \}, \{ \text{source}(3)=1, \text{cible}(3)=2, \text{étiquette}(1)=\text{"dormir"}, \text{étiquette}(3)=\text{"1"} \} \rangle$. Ce type de représentation n'étant pas particulièrement lisible, nous lui préférons une représentation graphique. La figure suivante encode la même information. Les objets y sont représentés par des formes géométriques (la forme indiquant le type de l'objet) et les fonctions sont représentées par des arcs en pointillés orientés de l'argument vers la valeur.

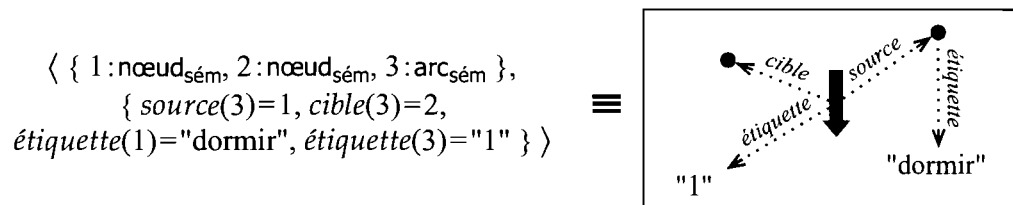


Figure 31 — Une structure GUST et sa représentation graphique explicite

Toutefois, cette structure n'est pas tout à fait complète. En effet, les structures utilisées en GUST doivent être polarisées puisque c'est cette polarisation qui contrôle la combinaison des règles.

3 Le système des polarités

Une **structure polarisée** est une structure dont tous les objets sont associés par une fonction polarisante à une valeur appartenant à un ensemble fini \mathcal{P} de

7. Merci à Kim Gerdes pour avoir porté ce fait à notre attention.

polarités. Les polarités servent à contrôler la saturation des structures construites par les grammaires. À notre connaissance, le premier à avoir utilisé à cette fin la polarisation d'objets dans une grammaire était Nasr (1995). Par la suite, Duchier et Thater (1999) ont développé cette idée, et Kahane (2000a), Gerdes (2001) et Perrier (2002) ont proposé des formalismes différents basés sur le concept de polarité. Le formalisme que nous utilisons est essentiellement celui des grammaires d'unification polarisées (GUP) proposé par Kahane (2004). GUP est un formalisme générique, c'est-à-dire qu'il n'est pas associé à une théorie linguistique en particulier. Il a été utilisé pour la formalisation de GUST d'abord par Kahane & Lareau (2005a).

Nous utilisons un système à deux polarités, que nous appellerons respectivement *noire* et *blanche*: $\mathcal{P} = \{\blacksquare, \square\}$. La polarité noire représente la saturation alors que la blanche représente au contraire la non-saturation, en d'autres termes un «besoin», ou une «place ouverte»⁸. Ce dispositif permet d'indiquer, par exemple, dans la structure sémantique de '[X] dormir' donnée plus haut, que cette structure est incomplète puisque 'dormir' doit avoir un actant. On n'a alors qu'à polariser en blanc le nœud dépendant pour indiquer qu'il faut «comblé un besoin»:

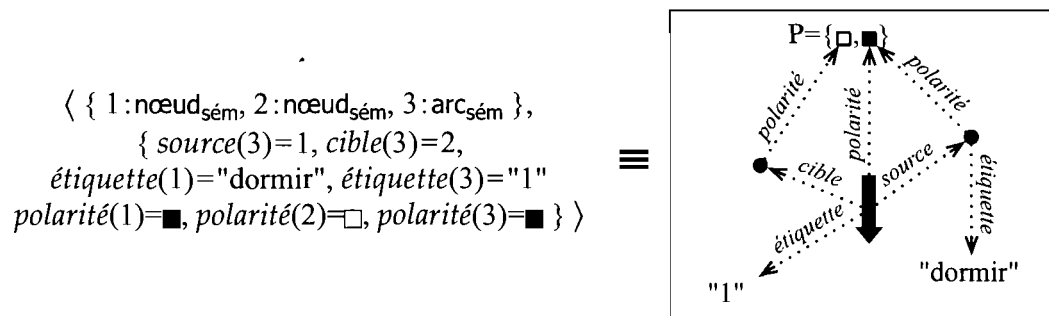


Figure 32 — Une structure polarisée et sa représentation graphique

On comprendra que même la représentation graphique devient vite illisible. Nous allons donc utiliser dorénavant des représentations où un certain nombre d'informations seront implicites. Notamment, les fonctions *source* et *cible* seront

8. Kahane & Lareau (2005a, 2005b) avaient une polarité supplémentaire grise, mais nous verrons qu'elle n'est pas nécessaire puisqu'elle équivaut formellement à une absence de polarité. Bien que du point de vue du linguiste qui développe la grammaire il peut être souhaitable d'avoir la polarité grise, nous avons décidé de ne pas l'utiliser dans cette thèse.

toujours omises dans les figures, l'orientation des arcs étant suffisante. D'autre part, les polarités seront indiquées directement sur les objets, qui seront selon le cas noirs ou blancs. Enfin, les étiquettes seront tout simplement placées à côté des objets, sans que la fonction d'étiquetage ne soit explicitement montrée. Ainsi, la structure de la Figure 32 sera représentée ainsi :

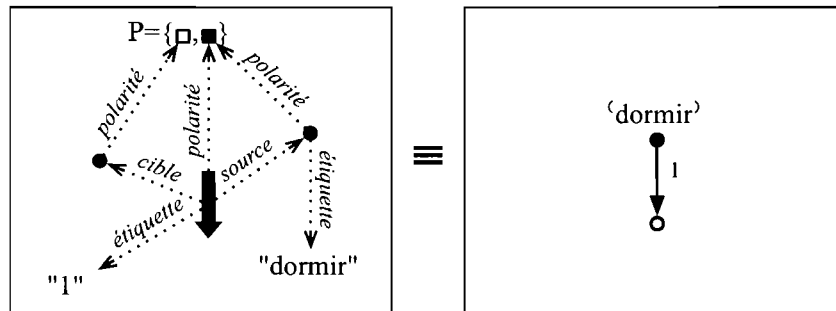


Figure 33 — Une représentation graphique simplifiée pour une structure polarisée

L'ensemble \mathcal{P} est également muni d'une opération (commutative et associative), le **produit**, que l'on note « \cdot ». Cette opération est nécessaire pour la combinaison de structures polarisées. En effet, lorsque deux objets polarisés a et b sont unifiés, ils donnent naissance à un nouvel objet c , dont la polarité est le produit des polarités de a et de b : $polarité(c) = polarité(a) \cdot polarité(b)$ (nous y reviendrons à la section suivante). Le produit des polarités est défini par le tableau suivant (où « \perp » représente l'échec de l'opération):

\cdot	□	■
□	□	■
■	■	\perp

Tableau XXV — Le produit des polarités de GUST

Il est intéressant de noter que cette opération induit sur l'ensemble \mathcal{P} un ordre $\square < \blacksquare$, de sorte que le produit de deux polarités donne toujours celle dont le rang est le plus élevé, c'est-à-dire que $p_i \cdot p_j = \max\{p_i, p_j\}$ ⁹. Ainsi, les objets, en s'uni-

9. Il existe aussi des systèmes où $p_i \cdot p_j \geq \max\{p_i, p_j\}$. Par exemple, Kahane (2004) utilise un système où le produit des polarités $+$ et $-$ est la polarité noire, dont le rang est plus élevé que celui de $+$ et de $-$.

fiant, tendent vers la polarité noire. En d'autres termes, les structures, en se combinant, seaturent.

En plus de l'ensemble \mathcal{P} , il faut également définir un sous-ensemble N de \mathcal{P} qui contient les **polarités neutres**. Il s'agit d'un sous-ensemble strict, c'est-à-dire que les polarités de \mathcal{P} ne peuvent pas toutes être neutres. Nous utilisons un système à une polarité neutre: $N = \{\blacksquare\}$. Cette polarité sert à mettre un terme au processus d'unification des structures (que nous décrivons en détail à la section suivante). Tant qu'une structure contient des objets non neutres, elle est en quelque sorte «active», à la manière d'un atome dont la valence n'est pas neutre. Quand, après s'être unifiée avec d'autres structures, tous les objets ont été neutralisés, le processus d'unification s'arrête naturellement, de la même façon que les atomes cessent d'attirer d'autres atomes quand leur valence a été saturée et qu'une molécule stable s'est formée. Nous allons maintenant voir comment se fait l'unification des structures polarisées. Mais d'abord, une dernière remarque s'impose au sujet des polarités.

Libre au linguiste de définir un autre système de polarités selon ses besoins. Cependant, certains systèmes ne sont pas appropriés pour la description des langues. Il faut restreindre la puissance du formalisme en imposant les contraintes suivantes sur le système de polarités:

- 1) Contrainte de **dynamisme**: N doit contenir au moins une polarité et être un sous-ensemble au sens strict de \mathcal{P} (de sorte qu'il existe au moins une polarité neutre et une non-neutre, ce qui permet leur interaction, assurant ainsi le dynamisme de la grammaire).
- 2) Contrainte de **monotonie**: \mathcal{P} doit être ordonné, et le produit de deux polarités doit toujours donner une polarité d'un rang supérieur ou égal au rang le plus élevé des deux (de sorte que le processus d'unification des structures ne puisse pas boucler à l'infini et qu'il progresse inexorablement vers les polarités les plus élevées dans l'échelle).

- 3) Contrainte de **finalité**: la ou les polarités ayant le plus haut rang dans \mathcal{P} doivent être neutres (de sorte que les structures tendent à se neutraliser).
- 4) Contrainte de **déclarativité**: le produit des polarités doit être commutatif et associatif (de sorte que les règles de la grammaire puissent se combiner dans n'importe quel ordre).

La dernière contrainte est plus faible que les autres. Si elle est violée, on peut obtenir une grammaire qui fonctionne, mais l'ordre d'application des règles sera pertinent, ce que nous ne souhaitons pas. Par contre, les autres contraintes doivent absolument être respectées, sans quoi la grammaire risque de ne pouvoir décrire aucun langage.

4 L'unification de structures polarisées

Les règles GUST (qui sont des structures polarisées) se combinent selon un processus bien précis. La **combinaison** de deux structures A et B, notée « \oplus », donne une nouvelle structure par l'unification d'au moins un objet de A et de B. La combinaison de structures repose donc entièrement sur l'opération d'**unification** d'objets, notée « \sqcup ». Le processus d'unification de deux objets implique un certain nombre d'opérations qu'il convient d'explicitier ici.

Pour unifier un objet a de la structure A et un objet b de B, la première condition à vérifier est que les deux objets soient du même type. Il n'est pas possible par exemple d'unifier un arc et un objet grammémique, pas plus qu'il n'est possible d'unifier un nœud sémantique avec un nœud syntaxique. L'objet c obtenu par unification de a et de b aura bien sûr le même type que ces derniers.

Ensuite, il faut vérifier que les fonctions liées à ces objets peuvent également s'unifier. Si une fonction f est associée à a mais pas à b , alors cette fonction est simplement transférée à l'objet c , c'est-à-dire qu'on ajoute à la nouvelle structure C à laquelle c appartient la fonction f telle que $f(c) = f(a)$ ¹⁰. Si par contre une même fonction f est associée, dans leur structure respective, à la fois à a et à b , alors il faut

que $f(a)$ et $f(b)$ puissent s'unifier, ce qui se vérifie différemment selon le type de valeur :

- Dans le cas des fonctions d'étiquetage, il s'agit simplement de vérifier que $f(a) = f(b)$. Si c'est bien le cas, alors l'objet c sera associé à la fonction f avec la même valeur : $f(c) = f(a) = f(b)$.
- Dans le cas des fonctions structurantes, les valeurs $f(a)$ et $f(b)$ sont des objets qu'il faut alors tenter d'unifier. Si cette unification échoue, celle de a et b échoue également. Si au contraire elle réussit, alors on associe à c la fonction f telle que $f(c) = f(a) \sqcup f(b)$.
- Finalement, dans le cas des fonctions de polarisation, il s'agit seulement de calculer la polarité de c à partir de celles de a et de b : $f(c) = f(a) \cdot f(b)$. Cette opération peut toutefois échouer si $f(a) \cdot f(b) = \perp$, ce qui fait échouer l'unification de a et de b . En d'autres termes, d'après le système de polarités que nous avons adopté, il ne sera jamais possible d'unifier deux objets de polarité noire.

Toutes les fonctions associées aux objets unifiés sont nécessairement identifiées, comme le sont les traits quand on unifie deux structures de traits dans d'autres formalismes. Tout comme dans ces formalismes d'ailleurs, la combinaison des structures est un processus récursif. Lorsqu'on combine deux structures, la structure qui en résulte peut à son tour être combinée à n'importe quelle autre structure, y compris celles qui ont déjà été utilisées pour la former.

10. Lors du transfert des fonctions structurantes, la valeur est un objet d'une des structures à combiner. Dans la structure résultante, il faut ajuster cette valeur pour qu'elle soit le nouvel objet qui a été créé (par unification ou par recopie) à partir de l'objet donné comme valeur dans la structure de départ.

À titre illustratif, voyons comment s'effectue la combinaison des structures A et B suivantes (que nous donnons sous forme explicite et sous forme graphique):

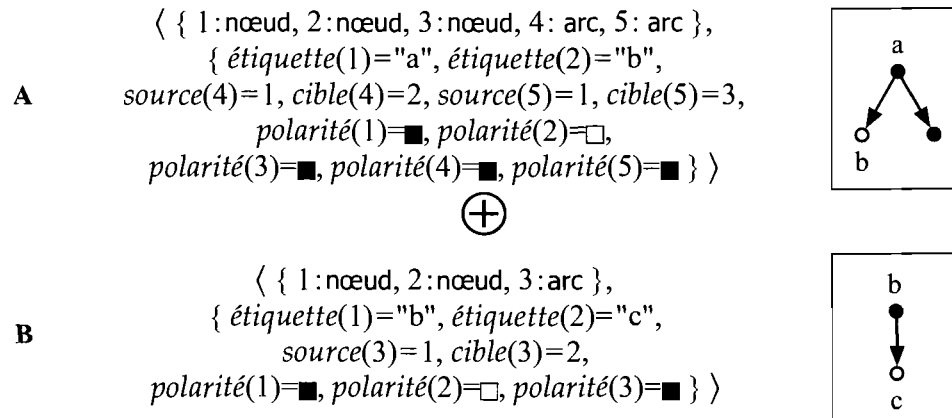


Figure 34 — Deux structures polarisées à combiner

Les objets déjà saturés ne peuvent pas s'unifier avec d'autres objets saturés. Ainsi, par exemple, le nœud 3 de A (celui sans étiquette) et le nœud 1 de B (celui étiqueté «b») ne peuvent pas s'unifier puisque le produit de leurs polarités échoue. Par ailleurs, les étiquettes associées aux objets contrôlent aussi la combinaison des structures. Ainsi, le nœud étiqueté «a» dans A ne peut pas s'unifier avec le nœud étiqueté «c» dans B, puisque leurs étiquettes ne peuvent pas être identifiées. Par contre, on peut unifier, par exemple, le nœud 3 de A (sans étiquette) avec le nœud 2 de B (étiqueté «c»). La polarité du nœud résultant sera alors $\square \cdot \blacksquare = \blacksquare$. En unifiant également les nœuds étiquetés «b», on obtient la structure C suivante:

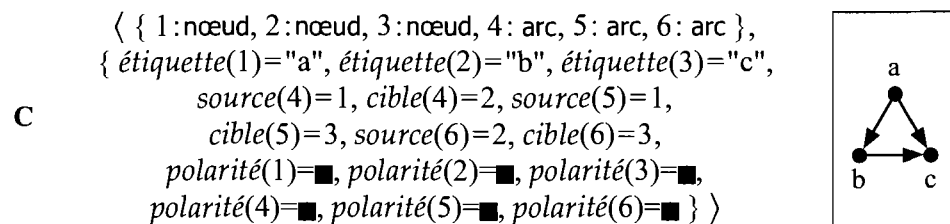


Figure 35 — Le résultat de la combinaison de deux structures polarisées

L'unification des objets n'est toutefois pas toujours obligatoire. Pour que deux structures se combinent, il suffit qu'un seul de leurs objets s'unifie. Lorsque toutes les unifications possibles pour deux structures sont effectuées, nous disons qu'il s'agit d'une **combinaison totale**. Lorsque, au contraire, on n'unifie qu'une

partie des objets qui pourraient s'unifier, on parle alors de **combinaison partielle**. Notons qu'il peut y avoir plus d'une combinaison totale possible pour une même paire de structures. Ce n'est pas le cas dans notre exemple, mais on a toutefois deux combinaisons partielles possibles, soit en n'unifiant pas les objets 2 de A et 1 de B, soit en ignorant la possibilité d'unification de 3 de A et 2 de B. On peut donc obtenir deux structures supplémentaires :

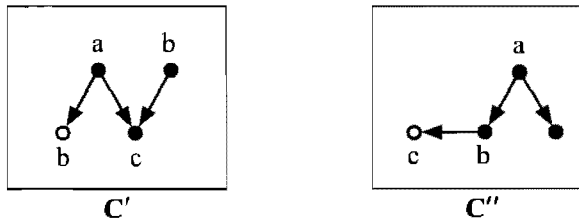


Figure 36 — La combinaison partielle de deux structures polarisées

Ceci conclut notre présentation du formalisme des GUP. Ce formalisme est utilisé dans tous les modules de GUST, et c'est là un des points forts de la théorie. Nous allons maintenant voir comment ces modules s'articulent pour former une grammaire cohérente.

5 Les modules de GUST et leur articulation

Nous avons vu plus haut (§ 1, p. 224) qu'un modèle GUST est composé de quatre grammaires de bonne formation qui construisent des représentations linguistiques valides et de trois grammaires d'interface qui assurent la mise en correspondance des structures construites par les grammaires de bonne formation. Les grammaires de bonne formation sont la grammaire sémantique ($\mathcal{G}_{\text{sém}}$), la grammaire syntaxique ($\mathcal{G}_{\text{synt}}$), la grammaire morphotopologique ($\mathcal{G}_{\text{topo}}$) et la grammaire phonologique ($\mathcal{G}_{\text{phon}}$). Les grammaires d'interface sont l'interface sémantique-syntaxe ($\mathcal{I}_{\text{sém-synt}}$), l'interface syntaxe-morphotopologie ($\mathcal{I}_{\text{synt-topo}}$) et l'interface morphotopologie-phonologie ($\mathcal{I}_{\text{topo-phon}}$). Toutes ces composantes forment la grammaire d'une langue. Nous utiliserons le terme «module» pour désigner tant les grammaires de bonne formation que les grammaires d'interface. Nous allons d'abord présenter chacun des modules en question, puis nous verrons com-

ment ils s'articulent. Cette section reprend l'essentiel des articles de Kahane & Lareau (2005a, 2005b) en y ajoutant une description complète des types d'objets.

5.1 Les grammaires de bonne formation

Les règles de bonne formation des structures de la TST sont très souvent implicites. En principe, elles doivent, par exemple, donner la liste de tous les sémantèmes pouvant apparaître dans une représentation sémantique dans une langue donnée (*cf.* Polguère 1990). Ces règles ne servent qu'à vérifier les structures données en entrée au modèle ou produites par celui-ci. Les grammaires de bonne formation de GUST peuvent fonctionner de deux façons: elles peuvent générer des structures (sémantiques, syntaxiques, morphotopologiques ou phonologiques) à partir de rien, ou encore elles peuvent vérifier la bonne formation d'une structure déjà existante. Nous verrons plus loin (§ 5.3.2, p. 256) que ce dernier mode de fonctionnement peut également entraîner la modification d'une structure incomplète de façon à obtenir une structure bien formée. Puisque la nature des représentations varie d'un niveau à l'autre, il va de soi que les structures manipulées par les différentes grammaires de bonne formation ainsi que les objets qui les composent ne seront pas du même type.

Dans sa première formulation de GUST, Kahane (2002) ne prévoyait pas de module distinct pour assurer la bonne formation des représentations linguistiques. Le principe à la base de sa théorie étant que chaque règle décrit un signe, il ne devait y avoir que des modules d'interface. Dans un article subséquent sur les signes grammaticaux dans l'interface sémantique-syntaxe de GUST, Kahane (2003b) fait mention de règles simples de bonne formation syntaxique, mais sans préciser où et dans quel format elles doivent être décrites. C'est lors de la reformulation de GUST en GUP par Kahane & Lareau (2005a, 2005b) qu'ont été introduits les modules de bonne formation sémantique et syntaxique. Cependant, nous n'avons pas discuté dans cet article de leur utilité, ni de ce qu'ils doivent contenir précisément. Si on suit l'esprit de Kahane (2002), les grammaires de bonne formation doivent être limitée à la description de phénomènes aussi généraux que possi-

ble. Tout ce qui est lié à un signe en particulier doit être décrit dans les grammaires d'interface.

Cependant, les correspondances entre niveaux adjacents ne sont pas bi-univoques. Un même objet peut donc apparaître dans plusieurs règles qui le mettent en correspondance avec divers objets de niveaux adjacents. Par exemple, le sémantème 'dormir' peut être lexicalisé par DORMIR, DODO, SOMMEIL, etc. Chaque lexicalisation est décrite par une règle différente. Mais il s'agit toujours du même sémantème, dont on peut vouloir, par exemple, typer les actants (le dormeur doit être vivant) ou fournir une décomposition sémantique. Alors, dans quelle règle faut-il le faire? Faut-il répéter cette information dans toutes les règles? Inversement, un même lexème peut être mis en correspondance avec différents sens s'il fait partie de phrasèmes. Par exemple, les locutions 'TIRER SON ÉPINGLE DU JEU', 'TIRER LE DIABLE PAR LA QUEUE' ou 'TIRER LES FICELLES' sont mises en correspondance avec leur sens respectif par autant de règles du module d'interface. Toutefois, elles sont toutes construites avec le même lexème TIRER, qui a des propriétés identiques dans toutes ces locutions: il s'agit d'un verbe, il doit avoir une finitude, peut-être un mode, un temps, etc. Faut-il répéter cette information dans toutes les règles de correspondance où TIRER apparaît? En décrivant la combinatoire sémantique et syntaxique des signes dans les règles de correspondance, comme le fait Kahane (2002), on est forcé d'introduire une certaine redondance. D'un point de vue théorique, il n'y a pas vraiment de problème; mais d'un point de vue pratique, le risque d'erreur de la part du linguiste augmente considérablement si l'information grammaticale est encodée de manière diffuse.

Pour cette raison, nous préférons décrire une partie de la combinatoire des signes dans les grammaires de bonne formation sémantique et syntaxique. Cette information est en fait une projection des niveaux plus superficiels que la syntaxe (ou plus profonds que la sémantique — mais on sort alors de la langue).

5.1.1 La grammaire sémantique

La grammaire de bonne formation sémantique ($G_{\text{sém}}$) construit des graphes dont les nœuds représentent des sémantèmes et les arcs représentent les relations prédicat-argument entre ces sémantèmes. Ce faisant, elle doit s'assurer que chaque prédicat possède tous les arguments qu'il lui faut, c'est-à-dire que sa valence sémantique active est saturée.

Les nœuds sont des objets du type $\text{nœud}_{\text{sém}}$. Les objets de ce type peuvent porter une fonction *étiquette* qui retourne comme valeur un symbole de l'ensemble S , qui contient tous les sémantèmes de la langue décrite (par exemple, pour le français, $S = \{\text{'chien'}, \text{'perdre la tête'}, \text{'temps'}, \text{'paranoïa'}, \dots\}$).

Les arcs, quant à eux, sont du type $\text{arc}_{\text{sém}}$. Tous les objets de ce type doivent obligatoirement porter une fonction *source* et une fonction *cible*, qui retournent toutes deux un objet de type $\text{nœud}_{\text{sém}}$. Ils peuvent également porter une fonction *étiquette* qui retourne comme valeur un symbole de l'ensemble $\mathcal{R}_{\text{sém}}$, qui contient toutes les relations sémantiques possibles (pour toutes les langues, $\mathcal{R}_{\text{sém}} = \{1, 2, 3, 4, 5, 6\}$).

Tant les objets de type $\text{nœud}_{\text{sém}}$ que $\text{arc}_{\text{sém}}$ doivent également porter la fonction de polarité $p_{\text{sém}}$, qui retourne (puisque c'est une fonction de polarité) un des éléments de l'ensemble \mathcal{P} (cf. § 3, p. 229). Cette polarité est propre aux objets de la grammaire sémantique et sert à vérifier la saturation de la valence sémantique active des sémantèmes décrits par cette grammaire. Dans chaque règle de bonne formation sémantique, les objets polarisés en noir sont ceux réellement construits par la règle, alors que ceux polarisés en blanc représentent la valence sémantique des sémantèmes. La Figure 37 montre un fragment de la grammaire sémantique du français. Il s'agit de quatre règles qui indiquent que 'manger' est un prédicat à deux arguments (c'est-à-dire que la structure sémantique n'est pas complète si ses deux arguments ne sont pas instanciés, d'où la saturation de ses dépendants en blanc), que 'Pierre' et 'pomme' sont des noms sémantiques et que 'deux' est un prédicat à un seul argument. La figure montre également une structure qui peut être générée

par la combinaison de ces quatre règles. Cette structure est reconnue comme valide puisque tous ses objets sont saturés.

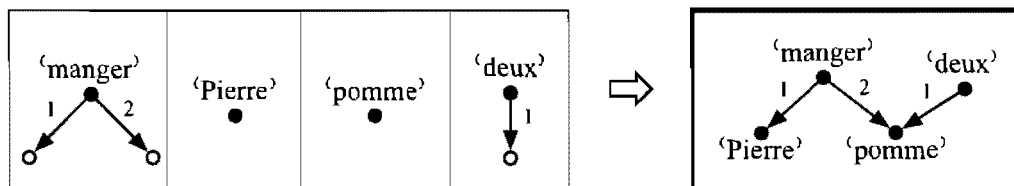


Figure 37 — Une grammaire sémantique générant une structure bien formée

Toutefois, rien n'empêche de combiner ces règles autrement et de créer des représentations qui sont bien saturées mais qui sont tout de même mal formées, comme le montre la Figure 38 :

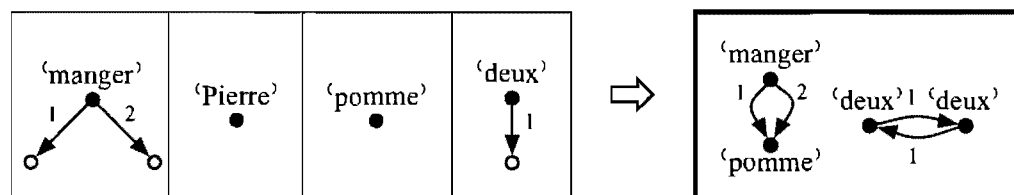


Figure 38 — Une grammaire sémantique générant une structure mal formée

La simple saturation de la valence des prédicats ne suffit pas à vérifier la bonne formation des structures sémantiques. Il y a deux autres propriétés qu'il faut vérifier au niveau sémantique. D'abord, il y a une contrainte structurelle générale, commune à toutes les langues : la structure doit être un graphe connexe et acyclique. Ensuite, il faut que les arguments d'un sémantème soient d'un type approprié. Par exemple, une activité, comme 'marcher', ne peut pas être le premier argument de 'regarder'.

La contrainte structurelle ne peut pas être vérifiée directement. En effet, le mécanisme des polarités ne peut vérifier que des propriétés locales (par exemple, si tel nœud a ou non un gouverneur), mais pour vérifier l'acyclicité d'un réseau il faut pouvoir vérifier qu'un nœud ne se gouverne pas lui-même, même indirectement. Par contre, la bonne formation structurelle d'un réseau peut se vérifier indirectement par sa mise en correspondance avec un arbre, puisqu'il n'est possible de construire un arbre dont les éléments sont en correspondance avec ceux d'un réseau si ce réseau n'est pas cyclique. Or, cette correspondance est vérifiée par la

grammaire d'interface sémantique-syntaxe, que nous présenterons plus loin (§ 5.3.2, p. 256).

Le problème du typage des actants, quant à lui, peut être résolu simplement en associant à chaque sémantème un attribut qui spécifie son «type» sémantique (sa composante sémantique centrale, c'est-à-dire, en général, son genre prochain, mais on peut choisir un typage plus grossier et utiliser le sémantème correspondant à un hyperonyme indirect). Il devient alors facile de vérifier le type sémantique d'un dépendant. Cette méthode est largement utilisée notamment dans les formalismes basés sur l'unification de structures de traits comme HPSG ou LFG. Elle peut aussi être implémentée en GUST, comme le montre la figure suivante, où on peut voir un typage des actants de 'regarder' sous forme d'une structure de traits et son équivalent en GUST :

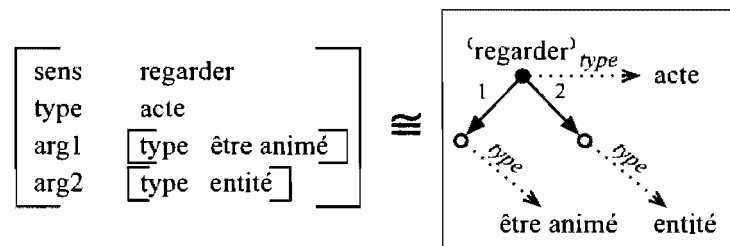


Figure 39 — Une méthode simple pour le typage des actants

Le problème avec cette solution est qu'elle présuppose une procédure implicite qui gère l'unification des types avec leurs sous-types. Par exemple, si on veut unifier le deuxième argument de 'regarder' avec la structure qui décrit le sémantème 'père', on doit faire appel à une procédure qui vérifie que le type de 'père' («individu») est bien un sous-type de «être animé». Cette procédure utilise une structure de données¹¹ additionnelle qui encode la hiérarchie des types (cf. Figure 40, ci-dessous).

11. Il s'agit en général d'ontologies (voir Psyché 2003 pour un état de l'art complet sur les ontologies). Dans le cadre de la TST cependant on utilise une hiérarchie d'étiquettes sémantiques (Milićević 1997, Polguère 2003a) où plutôt que d'organiser les choses du monde ce sont des sens d'une langue donnée que l'on organise.

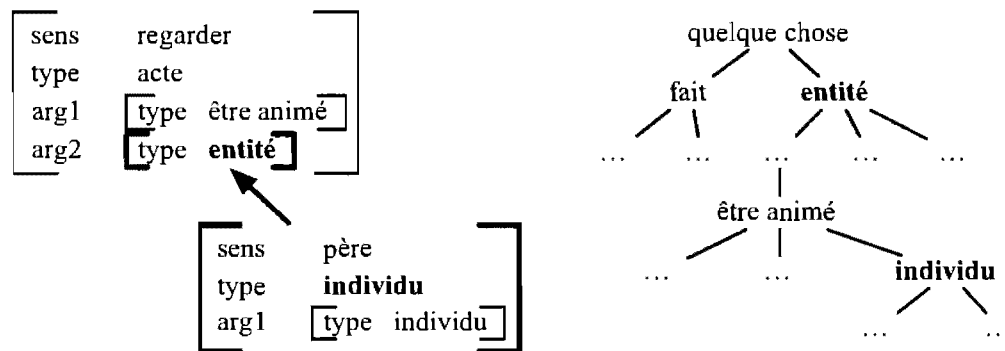


Figure 40 — Une procédure implicite pour l'unification de types et sous-types

Or, nous voulons justement décrire les phénomènes linguistiques en n'utilisant que la seule procédure d'unification polarisée présentée plus haut (§ 4, p. 233). L'appel à une procédure particulière pour assurer l'unification des types sémantiques s'insère donc mal dans notre modèle. Nous proposons plus loin (§ 6, p. 267) une solution qui permet de traiter le typage des actants sans avoir recours à une procédure particulière.

Voilà qui termine notre présentation de la grammaire sémantique. Ce type de grammaire ne construit que ce que la TST appelle la structure sémantique, c'est-à-dire un graphe de relations prédicat-argument entre des sémantèmes. Cette structure n'est que le squelette d'une représentation sémantique complète. On peut lui superposer d'autres structures, par exemple une structure communicative¹² (à ce sujet, voir notamment Mel'čuk 2001c). Nous laissons cependant de côté celle-ci, puisqu'elle n'est pas pertinente pour notre travail. Kahane (2005) propose également de superposer à la structure sémantique une structure encodant les relations de portée pour les quantificateurs dans le cadre de GUST. Nous laissons également de côté cette dernière pour les mêmes raisons. Nous proposons plus loin (§ 6, p. 267) une nouvelle structure pour représenter les décompositions sémantiques.

5.1.2 La grammaire syntaxique

La grammaire de bonne formation syntaxique (G_{synt}) construit des arbres de dépendance non ordonnés dont les nœuds représentent des lexèmes¹³. Ils sont

12. Qu'on appelle également *structure informationnelle* dans d'autres cadres théoriques.

reliés par des arcs qui représentent les relations syntaxiques entre ces lexèmes. Ils sont également associés à des objets grammémiques, qui représentent les grammèmes superficiels portés par les lexèmes. La grammaire syntaxique doit s'assurer que chaque lexème possède tous les grammèmes superficiels qui lui sont nécessaires selon sa partie du discours (par exemple, un nom doit porter un nombre, un verbe doit avoir un mode, etc.) et que les relations syntaxiques qui sont construites respectent certaines contraintes (par exemple, la relation sujet ne peut avoir comme gouverneur qu'un verbe fini).

Les nœuds de G_{synt} sont du type $\text{nœud}_{\text{synt}}$. Ces objets peuvent être associés à une fonction *étiquette*, qui retourne comme valeur un symbole de l'ensemble L , qui contient tous les lexèmes de la langue (par exemple, pour le français, $L = \{\text{CHIEN, PERDRE, LA, TÊTE, TEMPS, PARANOÏA, ...}\}$). Ils doivent aussi être liés à une partie du discours par la fonction *pdd*, qui retourne comme valeur un symbole de l'ensemble C , qui contient toutes les parties du discours pour la langue en question (pour le français, $C = \{\text{verbe, nom commun, article, pronom personnel, ...}\}$). Ils peuvent aussi être liés par diverses fonctions à des étiquettes qui représentent des propriétés intrinsèques du syntactique des lexèmes qu'ils représentent (par exemple, les noms en français porteront une fonction *genre* qui retournera comme valeur un des éléments de l'ensemble $\{\text{masculin, féminin}\}$).

Les nœuds syntaxiques peuvent également être liés à des objets grammémiques par des fonctions structurantes qui retournent comme valeur un objet de type *obj-gramm*. Ces fonctions portent le nom d'une catégorie flexionnelle de la langue décrite. Elles varient donc d'une langue à l'autre. Par exemple, pour le français, il existe les fonctions *nombre, mode, voix*, etc., puisque le nombre, le mode et la voix sont des catégories flexionnelles en français. Ces fonctions associent un lexème à un grammème d'une catégorie flexionnelle donnée.

13. Dans le cadre de GUST, il n'y a pas de niveau syntaxique profond. Les phrasèmes apparaissent donc au niveau syntaxique comme une configuration de lexèmes et non comme des nœuds simples. Kahane (2002, 2003c) discute en profondeur du statut de la représentation syntaxique profonde dans le cadre de GUST.

Les grammèmes sont représentés par des objets de type *obj-gramm*. Ceux-ci sont obligatoirement liés par la fonction *hôte* à l'objet de type *nœud_{synt}* qui les porte¹⁴. Ils peuvent être liés par la fonction *étiquette* à un symbole de l'ensemble \mathcal{G} , qui contient tous les grammèmes possibles pour la langue en question (par exemple, pour le français, $\mathcal{G} = \{\text{pluriel, présent, indicatif, féminin, ...}\}$). Nous insistons sur le fait que les grammèmes sont représentés par des *objets*, et non par des structures de traits associées aux nœuds syntaxiques. Par souci de lisibilité, nous les représentons toutefois comme des structures de traits dans nos figures, mais ce n'est qu'une métaphore visuelle. Ainsi, nous utilisons dans nos représentations graphiques les noms des fonctions comme noms d'attributs, et les étiquettes des objets de type *obj-gramm* comme valeur de ces attributs. Entre les deux cependant nous indiquons la polarité de l'objet *obj-gramm* par un losange. Dans ces «fausses structures de traits», nous mettons également les informations sur les propriétés intrinsèques du syntactique des lexèmes : leur partie du discours, leur genre, leur groupe de conjugaison, etc. Ces derniers ne sont toutefois pas représentés formellement par des objets mais simplement par une fonction qui retourne une étiquette, c'est pourquoi nous ne mettons pas de losange dans nos figures. Ainsi, la structure suivante se représente graphiquement comme suit :

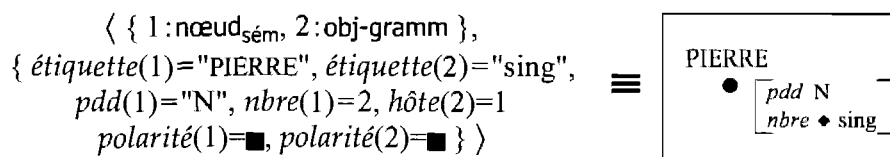


Figure 41 — La représentation d'un nœud syntaxique et de l'information grammaticale qui lui est rattachée

Nous insistons encore une fois sur la différence entre la représentation des traits de syntactique des lexèmes et celle des grammèmes. Les premiers sont représentés par des fonctions d'étiquetage qui retournent une simple étiquette (qui ne peut pas être polarisée), alors que les seconds le sont par des fonctions structurantes qui pointent vers un objet (lequel est polarisé et étiqueté).

14. Cette fonction permet de mieux contrôler l'unification des objets en évitant que plusieurs nœuds «partagent» le même objet grammémique.

Enfin, les arcs de $\mathcal{G}_{\text{synt}}$ sont des objets du type arc_{synt} . Tous ces objets doivent obligatoirement porter une fonction *source* et une fonction *cible*, qui retournent toutes deux un objet de type $\text{nœud}_{\text{synt}}$. Ils peuvent également porter une fonction *étiquette* qui retourne comme valeur un symbole de l'ensemble $\mathcal{R}_{\text{synt}}$, qui contient toutes les relations syntaxiques possibles. Ces relations varient d'une langue à l'autre (par exemple, pour le français, $\mathcal{R}_{\text{synt}} = \{\text{sujet, objet direct, circonstanciel, ...}\}$)¹⁵.

Tous les objets du module syntaxique doivent également porter la fonction de polarité p_{synt} , qui retourne un des éléments de l'ensemble \mathcal{P} . Cette polarité est propre aux objets de la grammaire syntaxique et sert à vérifier leur bonne formation en s'assurant que les fonctions associées à chacun sont bien instanciées: *source* et *cible* pour les arcs, *étiquette* et *pdd* pour les nœuds, ainsi que les fonctions les liant à des objets grammémiques, selon la partie du discours du nœud hôte (*nombre* pour les noms, *genre* et *nombre* pour les adjectifs, etc.).

La Figure 42 ci-dessous présente un fragment simplifié de la grammaire syntaxique du français ainsi qu'une structure qui peut être créée par ces règles.

N.B.: Afin de simplifier les figures dans les pages qui suivent, nous utilisons un système flexionnel simplifié pour les verbes où nous faisons abstraction des distinctions mode ~ finitude et temps ~ décalage (cf. Ch. VI, § 1, p. 116 et Ch. VI, § 6, p. 177). Les structures données en exemple tout au long du présent chapitre ne servent qu'à illustrer le formalisme et *ne font pas partie* de notre modèle formel de la flexion verbale en français, qui sera présenté au Chapitre VIII.

On peut distinguer dans cette grammaire quatre types de règles:

- La toute première règle est une structure initiale. Elle décrit un sommet syntaxique valide pour le français (il peut exister d'autres structures initiales alternatives, par exemple pour des phrases à l'impératif ou encore des phrases averbales¹⁶). Cette règle indique qu'un verbe à l'indicatif peut être

15. Mel'čuk & Pertsov (1987) et Iordanskaja & Mel'čuk (à paraître) proposent une méthodologie pour établir l'inventaire des relations syntaxiques (de surface) d'une langue et dressent la liste de celles, respectivement, de l'anglais et du français.

16. Voir notamment Lefevre (1999) au sujet des phrases averbales.

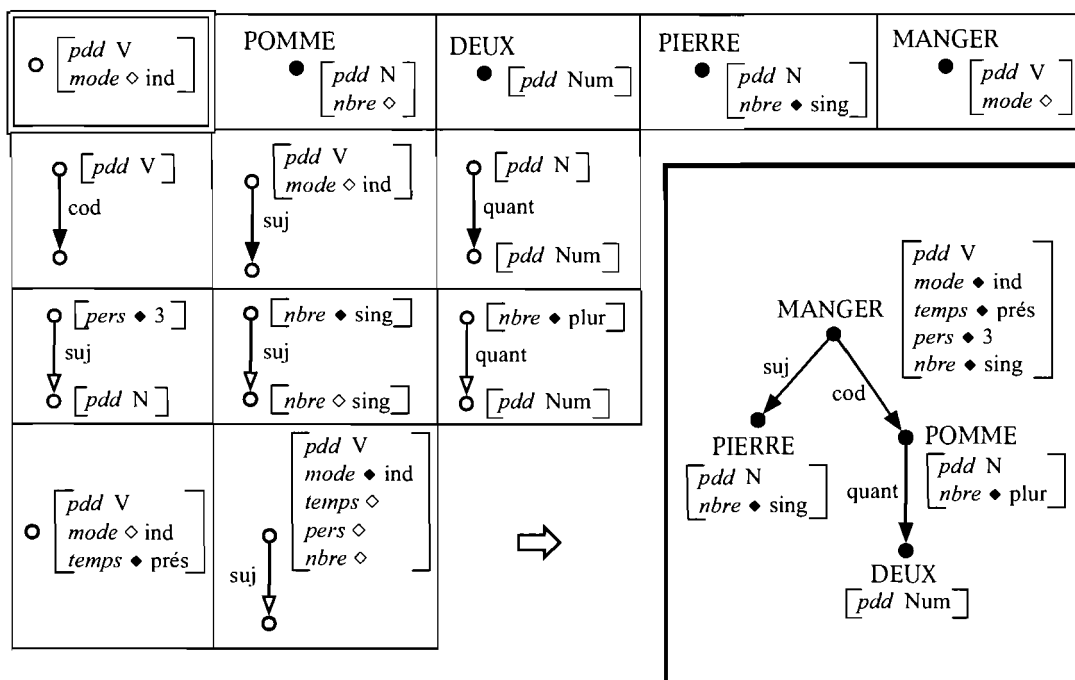


Figure 42 — Une grammaire syntaxique générant une structure bien formée

le sommet syntaxique d'une phrase¹⁷. Elle ne construit rien, mais sert plutôt à imposer des contraintes sur les lexèmes qui peuvent occuper cette position dans la représentation syntaxique.

- Nous avons ensuite quatre **règles lexicales**. Les règles lexicales ne font qu'énumérer l'ensemble des lexèmes de la langue, avec pour chacun sa partie du discours ainsi que les grammèmes superficiels qui lui sont nécessaires. Par exemple, POMME, en tant que nom, doit avoir un nombre (ce qui est exprimé par la polarisation blanche de l'objet grammémique). Alors que dans $G_{\text{sém}}$ on donnait pour chaque sémantème sa valence active, il n'est pas nécessaire de donner la valence syntaxique des lexèmes dans G_{synt} . Celle-ci est vérifiée par le module d'interface sémantique-syntaxe (§ 5.3, p. 254).
- Viennent ensuite, à la seconde ligne du tableau, trois **règles sagittales**. Ces règles décrivent l'ensemble des relations syntaxiques possibles dans la

17. En fait, nous avons vu au Chapitre VI que nous distinguons deux catégories flexionnelles, finitude et mode, là où la plupart des grammaires ne voient que le mode (cf. Ch. VI, § 1, p. 116, et Ch. VI, § 3, p. 134). Il serait donc plus exact de dire que c'est un verbe fini qui peut être le sommet syntaxique, mais dans le but de simplifier notre exposé du formalisme, nous préférons nous en tenir pour l'instant à un modèle plus proche des grammaires traditionnelles.

langue en question. Il est également possible de spécifier la partie du discours des nœuds liés par une relation donnée, ou même d'exiger la présence de certains grammèmes. Par exemple, la seconde règle sagittale indique qu'un verbe peut avoir un sujet s'il est à l'indicatif (toutefois, rien n'empêche d'avoir une autre règle qui autorise aussi la relation sujet, par exemple, pour les verbes au subjonctif).

- Enfin, il y a des **règles grammémiques**. Ces règles ne font que donner l'inventaire des grammèmes possibles pour chaque catégorie flexionnelle ainsi que leur combinatoire. Par exemple, la dernière règle signifie que l'indicatif est un grammème de mode valide en français, et que tout verbe à l'indicatif doit porter également un grammème de temps, de personne et de nombre, et doit avoir un sujet. Dans cet exemple, nous traitons l'accord dans $\mathcal{G}_{\text{synt}}$, mais nous verrons une autre solution au Chapitre VIII (§ 2, p. 298 et § 3, p. 316).

Nous annonçons au début de ce chapitre une vision «atomiste» de la grammaire, où les phénomènes sont éclatés au maximum. Le fragment de grammaire syntaxique de la Figure 42 l'illustre bien. Il est intéressant de comparer cette grammaire avec des grammaires entièrement lexicalisées comme TAG, qui a fortement inspiré GUST¹⁸. En TAG, pour construire une représentation à peu près équivalente à celle de la Figure 42, quatre structures «élémentaires» suffisent : une qui construit MANGER et les arcs qui en dépendent (tout en faisant l'accord avec le sujet), une qui construit PIERRE, une autre DEUX et l'arc dont il dépend, et une qui construit POMME. En apparence, TAG est plus simple, puisqu'il faut moins de règles pour arriver au même résultat. Mais ce type de grammaire pose quelques problèmes. D'une part, par exemple, MANGER peut apparaître dans d'autres constructions (au passif, sans objet, sans sujet, etc.). Il faut donc avoir autant de règles qu'il y a de structures où peut apparaître ce verbe. D'autre part, d'autres verbes, par exemple, peuvent occuper la même position que MANGER. Encore une fois, il faut avoir autant de règles qu'il y a de tels verbes, et pour toutes les constructions

18. On pourra consulter Kahane (2002) pour une discussion plus détaillée.

possibles. Il en résulte une explosion combinatoire du nombre de règles, ce qui ne se produit pas quand on sépare les phénomènes. Ce qui est plus important, c'est que la description de phénomènes grammaticaux généraux se trouve dissoute dans la description du lexique. Dans sa version pure, TAG ne met pas bien en lumière des généralisations importantes. Une solution à ce problème, proposée par Candito (1999), a été le recours à une métagrammaire, c'est-à-dire une «grammaire de la grammaire», dont la sortie est une grammaire TAG, donc entièrement lexicalisée. Ce mécanisme permet en somme de décrire des phénomènes grammaticaux récurrents sans introduire de redondance (mais cette métagrammaire est toujours compilée en une grammaire lexicalisée). GUST présente les avantages d'une métagrammaire.

5.1.3 La grammaire d'arbres

La petite grammaire syntaxique présentée ici ne suffit pas à garantir la construction de représentations syntaxiques bien formées. Par exemple, ces mêmes règles peuvent se combiner autrement pour donner la structure suivante :

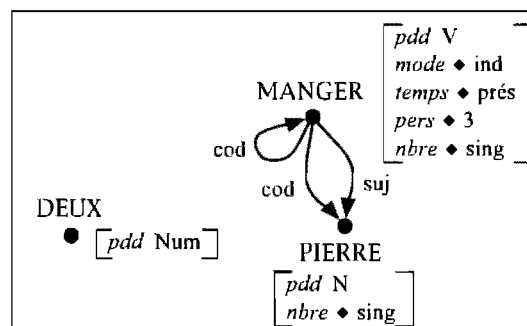


Figure 43 — Une structure syntaxique mal formée

Cette structure n'est pas valide pour plusieurs raisons, qui sont de deux types. D'une part certaines contraintes structurelles générales ont été violées (une représentation syntaxique doit être un arbre, ce qui n'est pas le cas ici). D'autre part la valence syntaxique active des lexèmes n'a pas été respectée (MANGER a deux compléments d'objet direct).

Alors qu'il n'était pas possible de vérifier avec une GUP qu'un réseau sémantique était structurellement bien formé, le formalisme permet de vérifier qu'une

structure est bien un arbre. Cela est possible puisque la bonne formation d'un arbre peut se vérifier par des contraintes locales sur les nœuds. En effet, si on impose comme contrainte sur les nœuds qu'ils soient tous gouvernés par un et un seul autre nœud, à l'exception du sommet qui ne doit pas être gouverné (décrit par une structure initiale), on s'assure de construire un arbre. Ainsi, la petite grammaire suivante, où un nœud n'est saturé que s'il est gouverné ou s'il s'agit de la structure initiale (en double encadré), ne peut construire que des arbres¹⁹.

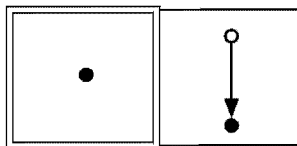


Figure 44 — Une grammaire d'arbres

Pour assurer la bonne formation des structures syntaxiques construites par $\mathcal{G}_{\text{synt}}$, il faut en quelque sorte coupler la grammaire donnée en Figure 42 à cette grammaire d'arbres. Pour ce faire, nous devons associer aux objets syntaxiques deux polarités. Puisque les polarités sont associées aux objets par des fonctions, rien n'empêche d'en assigner plus d'une à un même objet. Il suffit en effet d'utiliser une fonction différente pour chacune. Chaque polarité associée à un objet sert à vérifier une propriété différente. La polarité p_{synt} ²⁰, déjà introduite, servait à assurer la construction des objets en respectant leurs exigences syntaxiques. À celle-ci nous ajouterons une polarité p_{arbre} qui servira à vérifier que la structure construite est un arbre.

La représentation graphique de structures où les objets portent plusieurs polarités n'est pas toujours aisée. Partout où cela sera possible, nous masquerons les polarités qui ne sont pas pertinentes pour notre propos. Si toutefois il s'avère nécessaire de montrer dans une même figure plusieurs polarités, nous en choisirons une comme polarité principale, que nous indiquerons simplement par la couleur

19. Il s'agit de la grammaire proposée par Kahane (2004).

20. Par convention, nous faisons référence aux polarités portées par un objet par le nom de la fonction qui les associe à ce dernier. Il ne faut pas comprendre ici que p_{synt} est un élément de l'ensemble \mathcal{P} (cf. § 3, p. 229). Il s'agit d'une fonction qui retourne comme valeur un des éléments de cet ensemble. Lorsque nous écrivons «la polarité p_{synt} », c'est un raccourci pour «la polarité associée (à tel objet) par la fonction p_{synt} ».

des objets comme nous l'avons fait jusqu'ici, alors que les autres (qui nous dirons secondaires) seront données par de petits carrés juxtaposés aux objets (carrés dont la couleur indiquera la valeur de la polarité secondaire en question). Lorsque nous parlons de « polarité principale », il est important de souligner qu'il s'agit toujours de la polarité principale *d'une figure*. Une polarité n'est jamais une polarité principale *en soi*. On peut montrer la même structure de plusieurs points de vue, en choisissant chaque fois une polarité différente comme polarité principale, selon le propos que nous cherchons à illustrer. Par exemple, supposons une structure où les objets porteraient deux polarités, p_1 et p_2 . La figure suivante montre cette structure dans son encodage explicite ainsi que deux représentations graphiques possibles de celle-ci : dans la première, c'est p_1 qui est la polarité principale, alors que dans la seconde c'est p_2 .

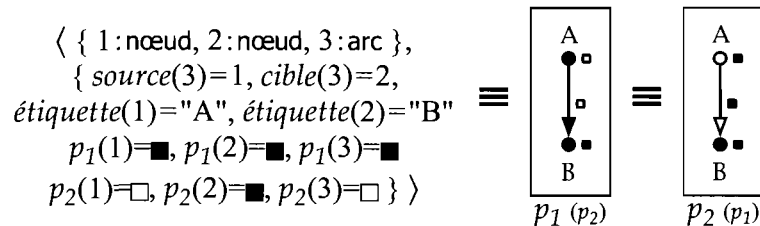


Figure 45 — La représentation graphique des structures multipolarisées

Maintenant, il s'agit de trouver la bonne façon de coupler les grammaires des Figures 42 et 44. La plus simple est d'ajouter aux structures de la Figure 42 une polarité p_{arbre} saturée pour tous les nœuds qui sont gouvernés (ainsi que pour la structure initiale) ou blanche pour les autres nœuds. On obtient alors la grammaire de la Figure 46 ci-dessous (p_{synt} est la polarité principale).

Cette solution manque toutefois d'élégance. Ce sont les mêmes règles qui décrivent à la fois les relations syntaxiques possibles dans la langue et les contraintes structurelles sur la représentation syntaxique. Le linguiste doit donc, quand il écrit une règle, penser à deux problèmes en même temps. Par ailleurs, il y a une certaine redondance dans cette grammaire. On voit bien que chaque fois qu'on construit un arc, on doit saturer en p_{arbre} le nœud qui en dépend. On aimerait donc séparer les deux problèmes et avoir deux jeux de règles qui se complètent : un qui décrit

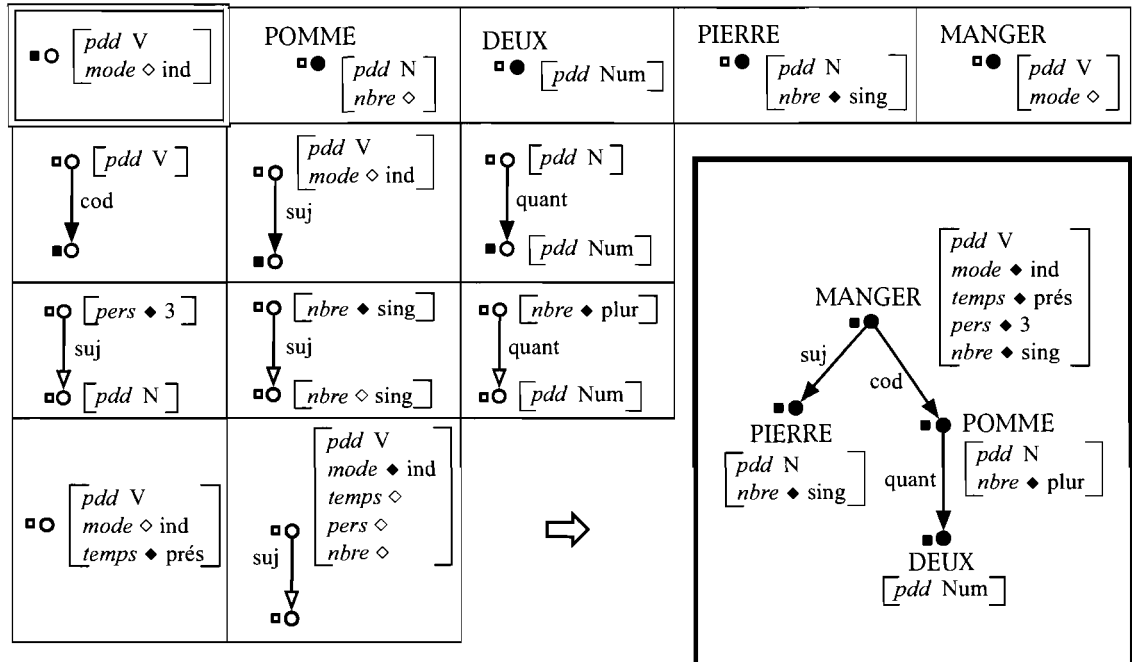


Figure 46 — Une grammaire syntaxique à deux polarités

l'ensemble des objets syntaxiques (ce sont les règles de la Figure 42) et l'autre qui contraint la structure à un arbre (ce sont les règles de la Figure 44). Le problème est de trouver comment articuler ces deux grammaires.

5.2 L'articulation des modules

L'**articulation** de deux grammaires se fait par un jeu de polarités. Supposons que nous ayons une grammaire \mathcal{A} dont les règles contiennent des objets saturés en p_a et une grammaire \mathcal{B} dont les règles contiennent des objets saturés en p_b . Pour articuler \mathcal{A} et \mathcal{B} , il faut faire en sorte que les objets construits par \mathcal{A} forcent l'application des règles de \mathcal{B} et vice-versa. Pour ce faire, on ajoute une polarité p_b blanche à tous les objets de \mathcal{A} et une polarité p_a blanche aux objets de \mathcal{B} . Cette polarisation supplémentaire ne concerne toutefois que les objets qui sont d'un type commun à \mathcal{A} et à \mathcal{B} (par exemple, si \mathcal{A} contient des objets de type t_1 et t_2 , et \mathcal{B} des objets de type t_2 et t_3 , seuls les objets de type t_2 , et tous ces objets, seront articulés). Nous dirons de ces objets qu'ils sont les **points d'articulation** de \mathcal{A} et \mathcal{B} . En général, les règles de \mathcal{A} ne doivent pas saturer la polarité p_b des objets et, inversement, les règles de \mathcal{B} ne satureront pas les objets en p_a . Ainsi, la seule façon de satu-

rer la polarité p_b blanche des objets construits par \mathcal{A} est d'appliquer les règles de \mathcal{B} (et vice-versa)²¹. Nous dirons que p_b dans la grammaire \mathcal{A} , et p_a dans \mathcal{B} , sont des **polarités d'articulation**, par opposition aux **polarités de construction**, qui sont celles effectivement saturées dans les règles.

Ce qui nous intéresse dans le cas présent est d'articuler la grammaire G_{synt} présentée à la Figure 42 et la grammaire d'arbres de la Figure 44 (que nous appellerons G_{arbre}). Les règles de G_{arbre} ne contiennent que des objets de type $\text{nœud}_{\text{synt}}$ et arc_{synt} , deux types qui se retrouvent également dans G_{synt} . Ce sont donc les objets de ces deux types qui serviront de points d'articulation. Tous les objets de G_{arbre} porteront une polarité p_{synt} blanche et tous les points d'articulation dans G_{synt} porteront une polarité p_{arbre} blanche. On obtient ainsi les deux grammaires ci-dessous, Figure 47 et Figure 48, p. 253 (nous montrons p_{synt} comme polarité principale dans les deux figures de façon à mettre en relief l'articulation).

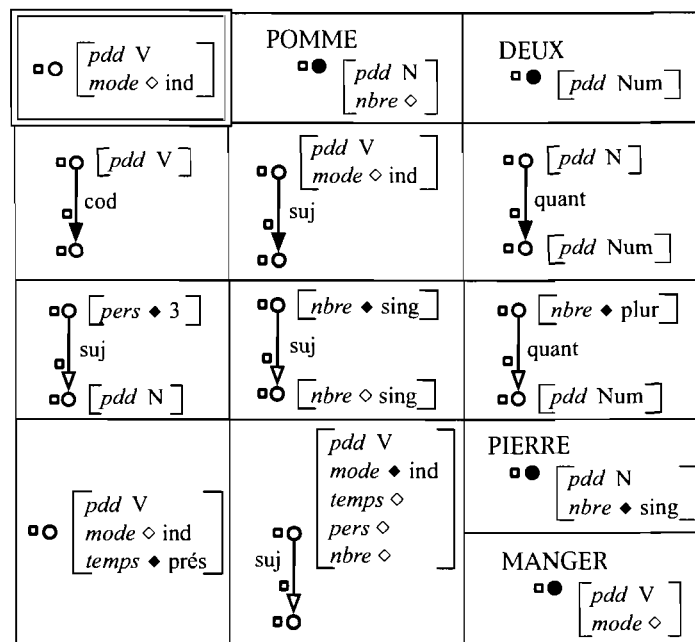


Figure 47 — Une grammaire syntaxique articulée

21. Nous verrons plus loin qu'une grammaire peut à l'occasion saturer la polarité d'articulation de certains objets pour bloquer leur mise en correspondance avec un niveau adjacent (cf. § 6, p. 267).

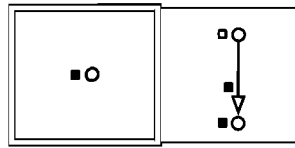


Figure 48 — Une grammaire d'arbres articulée

G_{synt} peut toujours construire des structures qui ne sont pas forcément des arbres, mais ses objets portent une polarité p_{arbre} blanche qui ne peut être neutralisée par G_{arbre} que si la structure en question est bien un arbre. La Figure 49 ci-dessous illustre la neutralisation mutuelle d'une structure construite par G_{synt} et d'un arbre construit par G_{arbre} .

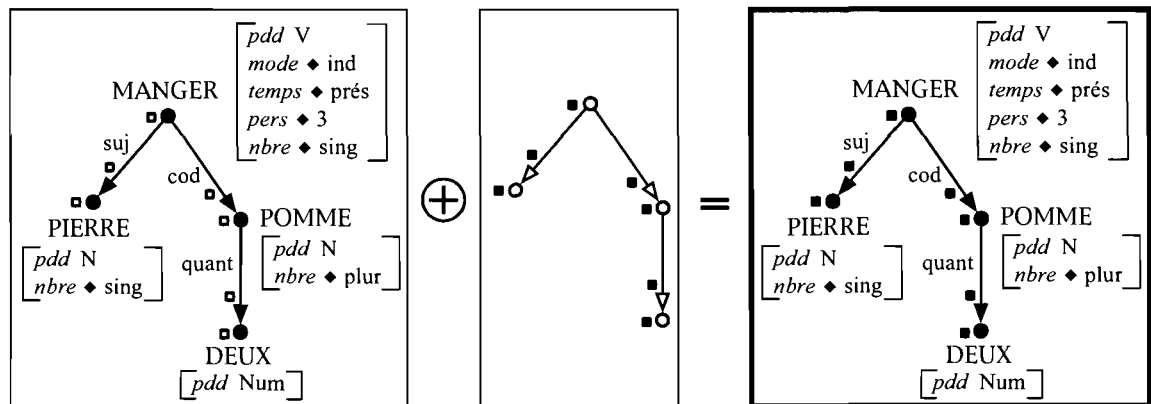


Figure 49 — Deux grammaires articulées produisant une structure syntaxique

Cette solution est plus élégante que la première puisqu'elle dégage une généralisation qui était cachée dans les règles de la Figure 46. D'ailleurs, les contraintes structurelles imposées par G_{arbre} seraient universelles selon cette théorie (si on accepte son postulat qui veut que l'organisation syntaxique de toutes les langues soit toujours décrite par un arbre de dépendance). Cette grammaire d'arbre pourrait donc en principe être réutilisée telle quelle pour n'importe quelle langue.

Strictement parlant, G_{arbre} constitue un des modules de GUST. Cependant, étant donné le nombre très limité de règles qu'il contient et son contact étroit avec G_{synt} , nous allons dorénavant le considérer comme un sous-module faisant partie de G_{synt} .

Voilà donc pour l'articulation des modules de GUST et les contraintes structurales sur les représentations syntaxiques. Nous avons également noté un autre problème dans la Figure 43 : le lexème MANGER pouvait avoir deux compléments d'objet direct, puisque G_{synt} ne décrit pas la valence syntaxique des lexèmes. Cette valence sera décrite par le module qui assure l'interface entre les représentations sémantiques et syntaxiques, module que nous allons maintenant présenter.

5.3 Les grammaires d'interface

Un modèle GUST comprend trois grammaires d'interface : l'interface sémantique-syntaxe, l'interface syntaxe-morphotopologie ainsi que l'interface morphotopologie-phonologie. Ces grammaires d'interface correspondent aux modules sémantique, syntaxique et morphologique de la TST. La première a été beaucoup plus étudiée que les deux autres dans le cadre de GUST. Une brève esquisse de l'interface syntaxe-morphotopologie a été présentée par Kahane & Lareau (2005a), mais il ne s'agit certainement pas de sa forme définitive. Les travaux actuels dans ce domaine portent plutôt sur la synchronisation d'une grammaire syntaxique avec une grammaire topologique à la Gerdes & Kahane (2001). Il existe déjà des descriptions de ce genre pour l'allemand (Gerdes 2002), le grec (Yoo 2003), le coréen (Yoo & Gerdes 2004), l'arabe (El-Kassas 2005) et le français (Gerdes & Kahane 2006). Toutefois, aucun de ces travaux n'utilise le formalisme des GUP. L'interface morphotopologie-phonologie quant à elle a toujours été le parent pauvre de la TST, et la situation n'est pas plus reluisante en GUST, même si les travaux de Yoo (2003) ont ouvert des pistes, notamment sur le calcul des constituants prosodiques. Nous ne savons pas comment encoder ce module dans le formalisme des GUP et nos connaissances limitées en phonologie ne nous permettent pas de remédier de façon satisfaisante à cette lacune. Dans le cadre de notre thèse, nous ne considérons que l'interface sémantique-syntaxe. Nous ne présenterons donc que cette dernière, non sans avoir d'abord introduit le concept de correspondance qui sous-tend toutes les grammaires d'interface. Cette section reprend les idées de Kahane (2000b) et Kahane & Lareau (2005a, 2005b).

5.3.1 *La notion de correspondance*

Deux structures A et B sont dites en **correspondance** lorsque les éléments qui les constituent sont liés par une relation sémiotique du type «signifié-signifiant». Depuis de Saussure, la correspondance entre un sémantème et la chaîne phonologique qui l'exprime est généralement considérée comme un signe linguistique²² et la plus petite unité de correspondance entre ces deux extrêmes est conçue comme un signe élémentaire. Pourtant, dans le cadre de GUST (et il en va de même pour la TST), même les signes élémentaires sont représentés non pas par une correspondance directe entre sémantème et chaîne phonologique, mais par un ensemble de correspondances plus simples encore (aux sémantèmes correspondent des objets syntaxiques, auxquels correspondent des objets morphologiques qui eux sont en correspondance avec une chaîne phonologique). C'est ce qui fait dire à Kahane (2001) que l'articulation du langage n'est pas double, mais qu'elle est encore plus complexe. Il appelle «signe» toute unité à deux faces, même si ces faces en questions sont des objets de représentations intermédiaires. C'est dans cette acception large que nous utiliserons aussi le terme «signe».

La correspondance entre les éléments de deux structures A et B est décrite par une **grammaire d'interface**²³. Le nombre de structures A et B possibles étant théoriquement infini, la grammaire ne peut évidemment pas établir la correspondance entre ces structures comme telles mais elle décrit plutôt, par le biais de correspondances élémentaires entre des objets, la correspondance entre les ensembles *A* et *B* auxquelles appartiennent les structures A et B. Ces ensembles ne sont en fait rien d'autre que les grammaires de bonne formation qui construisent A et B. La mise en correspondance des structures se fait donc par la synchronisation des grammaires de bonne formation (Kahane 2002).

Une grammaire d'interface peut être considérée selon le cas comme une grammaire équative, transductive ou générative (Kahane 2000b). Une grammaire équative prend en entrée deux structures et vérifie qu'elles sont bien en correspondance,

22. Mel'čuk (1993) apporte une troisième composante au signe : sa combinatoire.

23. On utilisera souvent le mot «interface» au lieu de l'expression «grammaire d'interface».

une grammaire transductive prend en entrée une structure A et construit la structure B qui lui correspond alors qu'une grammaire générative construit à partir de rien une paire de structures qui sont en correspondance. Un modèle GUST étant constitué d'une chaîne de grammaires d'interface, il peut fonctionner selon ces trois modes. GUST permet donc de vérifier qu'un sens et un texte donnés se correspondent, construire à partir d'une représentation sémantique le texte qui l'exprime (ou à l'inverse, à partir d'un texte, construire la représentation de son sens), ou encore générer des paires sens/texte qui sont en correspondance.

Ce qui nous intéresse particulièrement, c'est le fonctionnement transductif de GUST, puisque c'est celui qui modélise les processus linguistiques d'analyse et de synthèse. C'est toutefois dans un mode génératif que nous présenterons l'interface sémantique-syntaxe dans un premier temps pour des raisons de clarté, mais nous montrerons plus loin comment elle peut s'appliquer de manière transductive.

5.3.2 L'interface sémantique-syntaxe

La grammaire d'interface sémantique-syntaxe, que nous appellerons $I_{\text{sém-synt}}$, établit la correspondance entre l'ensemble de graphes sémantiques décrit par $G_{\text{sém}}$ et l'ensemble d'arbres syntaxiques décrit par G_{synt} . C'est donc elle qui assure, par exemple, la transition entre la structure sémantique de la Figure 37 (p. 240) et la structure syntaxique de la Figure 49 (p. 253), que nous reproduisons ici en masquant la polarité p_{arbre} .

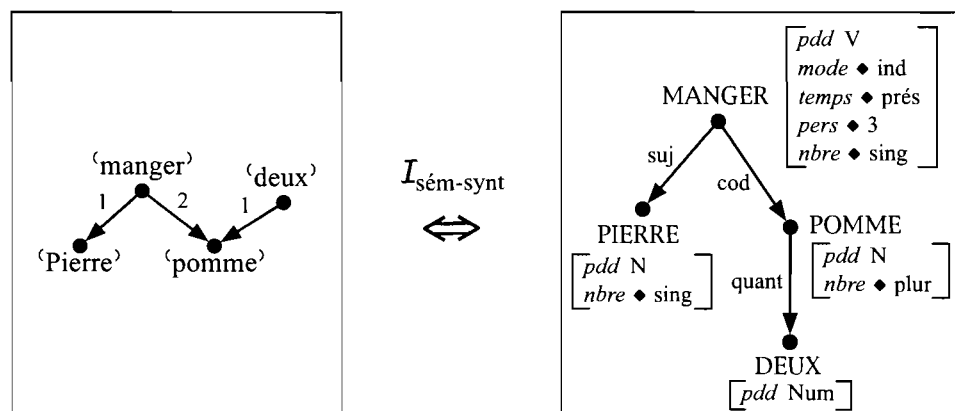


Figure 50 — La position de la grammaire d'interface sémantique-syntaxe

Cette mise en correspondance se fait par la construction d'un **graphe de correspondance**. Il s'agit d'un graphe complexe dans lequel on peut reconnaître deux sous-graphes, l'un étant une structure sémantique et l'autre étant une structure syntaxique. Les objets de ces deux sous-graphes sont du même type que dans les grammaires de bonne formation correspondantes mais ils ne portent pas les mêmes polarités que dans ces dernières. Les nœuds (et seulement les nœuds)²⁴ sémantiques et syntaxiques sont reliés par des liens de correspondance tels que définis plus haut (§ 5.3.1, p. 255), représentés par des objets de type correspondance. Ces objets servent à garder une trace des mises en correspondance effectuées, de façon à guider l'unification des structures. Ils sont associés à deux fonctions: *signifié*, qui retourne comme valeur un nœud du niveau sémantique, et *signifiant*, qui retourne comme valeur le nœud correspondant au niveau syntaxique. Dans les représentations graphiques toutefois, ces fonctions sont masquées et par convention tous les objets sémantiques sont placés à gauche et les objets syntaxiques à droite.

Tous les objets de $I_{\text{sém-synt}}$ doivent porter une polarité $p_{\text{sém-synt}}$. De façon générale, un objet sémantique ou syntaxique n'est saturé en $p_{\text{sém-synt}}$ que lorsqu'il est effectivement mis en correspondance avec un objet de niveau adjacent. Nous adoptons également comme principe que les objets de type correspondance ne doivent être saturés en $p_{\text{sém-synt}}$ dans une règle que lorsqu'ils lient des nœuds qui sont eux-mêmes saturés dans cette même règle (nous verrons plus loin pourquoi). La figure suivante montre un fragment de la grammaire d'interface sémantique-syntaxe du français. Les deux premières règles sont de simples règles de lexicalisation qui indiquent les lexèmes qui peuvent exprimer ces sémantèmes. Les deux autres règles ne sont en fait pas autre chose que des règles de lexicalisation non plus, mais comme il s'agit de prédicats, il faut en indiquer également la diathèse, c'est-à-dire

24. Nous aurions pu choisir de mettre en correspondance les arcs et les objets grammémiques. Cependant, les correspondances entre ces objets peuvent toujours être déduites puisqu'ils sont toujours liés à un nœud par une fonction (*source* ou *cible* pour un arc, *hôte* pour un objet grammémique).

la correspondance entre les actants sémantiques des sémantèmes et les actants syntaxiques des lexèmes qui les expriment.

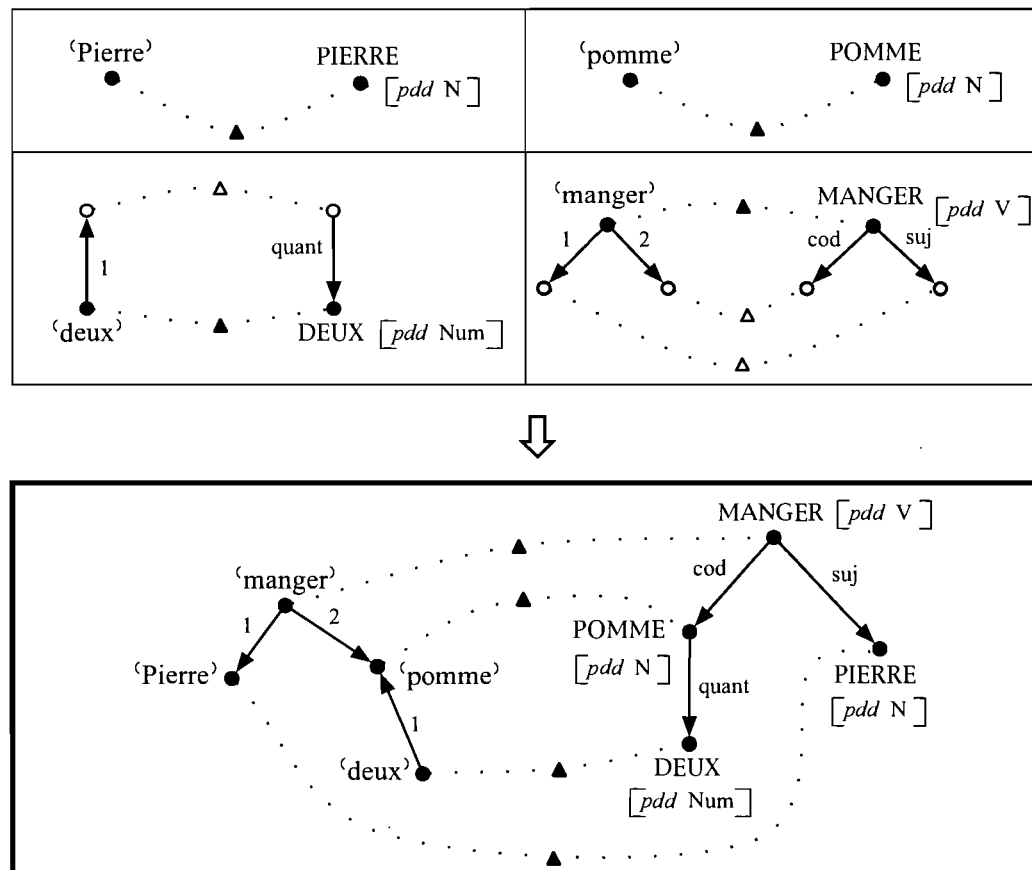


Figure 51 — Une grammaire d'interface sémantique-syntaxe générant une structure de correspondance

Les objets de type obj-gramm peuvent également être mis en correspondance avec les objets sémantiques qui encodent le sens du grammème en question (si ce dernier est sémantiquement plein, bien entendu; les grammèmes sémantiquement vides, comme les grammèmes d'accord, n'ont pas de correspondant au niveau sémantique). Par exemple, la règle de la Figure 52 ci-dessous met en correspondance un grammème de temps avec son sens (que nous ne décomposons pas pour l'instant — nous y reviendrons au Chapitre VIII).

La correspondance entre les nœuds de niveaux adjacents n'est pas toujours bi-univoque. D'abord, un nœud peut très bien n'être associé à aucun lien de correspondance. C'est le cas par exemple du sémantème 'présent' dans la règle ci-des-

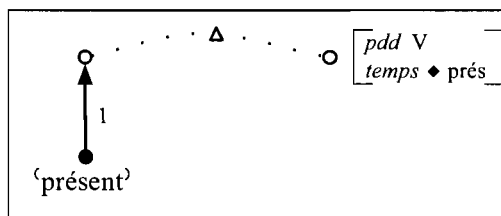


Figure 52 — La mise en correspondance d'un grammème et de son sens

sus²⁵. D'autre part, un même nœud peut avoir plusieurs correspondants. Ce n'est pas forcément le cas des phrasèmes, contrairement à ce que l'on serait porté à croire spontanément. En effet, pour la plupart des locutions tous les dépendants syntaxiques sont rattachés au même nœud, la tête syntaxique du phrasème²⁶. Il n'est donc pas nécessaire d'établir plusieurs liens de correspondance. Par contre, c'est toujours le cas lorsqu'on introduit des verbes support. Par exemple, la collocation *la crainte [de voir sa famille arrêtée] l'habite* peut se décrire par la règle suivante:

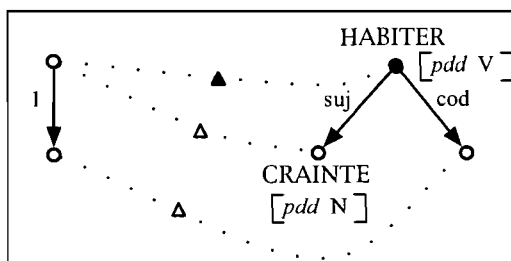


Figure 53 — Une correspondance multivoque

Le lexème CRAINTE conserve tous ses autres actants, qui doivent se rattacher sous lui, mais tout gouverneur doit se rattacher au verbe support. Le sémantème 'crainte' a donc deux correspondants au niveau syntaxique. Cet encodage modélise le fait que HABITER n'apporte pas de sens nouveau et que ce lexème sert avant tout à verbaliser CRAINTE et à le lier syntaxiquement à son premier actant sémantique.

Les structures de correspondance construites par $I_{\text{sém-synt}}$ peuvent être composées de graphes sémantiques ou d'arbres syntaxiques mal formés. Par exemple, la structure de correspondance de la Figure 51 est bien formée en tant que structure

25. Une correspondance implicite existe tout de même entre 'présent' et le grammème **présent**.

26. Ce n'est pas toujours le cas. Cf. par exemple, 'CASSER DU SUCRE SUR LE DOS'.

de correspondance (tous les objets sont saturés en $p_{sém-synt}$) mais l'arbre syntaxique qu'elle contient n'est pas valide en français puisque, par exemple, le verbe ne porte ni mode ni temps. Ce n'est pas le rôle de la grammaire d'interface de vérifier la bonne formation des structures sémantiques et syntaxiques qu'elle met en correspondance. Elle doit pouvoir accepter en entrée une structure mal formée et en fournir la structure de niveau adjacent correspondante, qu'elle soit aussi mal formée ou non. Toutefois, $I_{sém-synt}$ doit appeler les grammaires $G_{sém}$ et G_{synt} pour qu'elles valident ces structures. Cela se fait en articulant les trois grammaires. Nous avons vu lors de notre présentation de G_{synt} comment une grammaire d'arbre et une grammaire syntaxique pouvaient être articulées (cf. pp. 251 et ssq). Le principe est exactement le même ici, si ce n'est que dans le cas présent il y a deux articulations en jeu: l'articulation $G_{sém}-I_{sém-synt}$ et l'articulation $I_{sém-synt}-G_{synt}$.

Les points d'articulation $G_{sém}-I_{sém-synt}$ sont tous les objets d'un type commun à ces deux grammaires. Il s'agit des objets de type $nœud_{sém}$ et $arc_{sém}$. Puisque $G_{sém}$ et $I_{sém-synt}$ n'ont chacune qu'une seule polarité, ce sont évidemment ces polarités qui serviront à l'articulation. Ainsi, tous les objets de $G_{sém}$ porteront, en plus de leur polarité propre, une polarité $p_{sém-synt}$ blanche. Réciproquement, tous les objets de type $nœud_{sém}$ et $arc_{sém}$ dans $I_{sém-synt}$ porteront une polarité $p_{sém}$ blanche.

L'articulation $I_{sém-synt}-G_{synt}$, quant à elle, prend appui sur les objets de type $nœud_{synt}$, arc_{synt} ou $obj-gramm$. Tous ces objets devront porter dans $I_{sém-synt}$ une polarité p_{synt} blanche. Quant aux objets de G_{synt} , ils devront être polarisés en blanc pour la polarité propre à $I_{sém-synt}$, c'est-à-dire $p_{sém-synt}$. On pourrait se demander pourquoi ne pas articuler $I_{sém-synt}$ avec la grammaire d'arbre G_{arbre} plutôt qu'avec G_{synt} en utilisant p_{arbre} comme polarité d'articulation. La raison de ce choix est que G_{arbre} ne contient pas d'objets de type $obj-gramm$. Il serait donc impossible de mettre les grammèmes en correspondance avec les sens qu'ils expriment. Puisque G_{synt} et G_{arbre} sont déjà articulées, il n'est pas nécessaire d'articuler cette dernière avec l'interface sémantique-syntaxe, pas plus qu'il n'est nécessaire d'articuler directement $G_{sém}$ et G_{synt} puisqu'elles sont déjà liées indirectement par $I_{sém-synt}$.

Nous reprenons ci-dessous les trois fragments de grammaires présentés plus haut, cette fois avec les polarités d'articulation pertinentes pour l'interface sémantique-syntaxe. D'abord $G_{sém}$, avec $p_{sém}$ comme polarité principale et $p_{sém-synt}$ en polarité secondaire. Nous y ajoutons une règle pour le sens 'présent'.

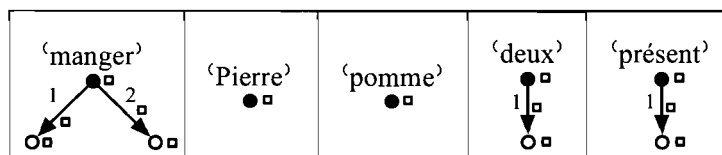


Figure 54 — Une grammaire sémantique articulée

Ensuite, $I_{sém-synt}$ avec comme polarité principale $p_{sém-synt}$ et en polarités secondaires $p_{sém}$ (à gauche des éléments sémantiques) et p_{synt} (à droite des éléments syntaxiques).

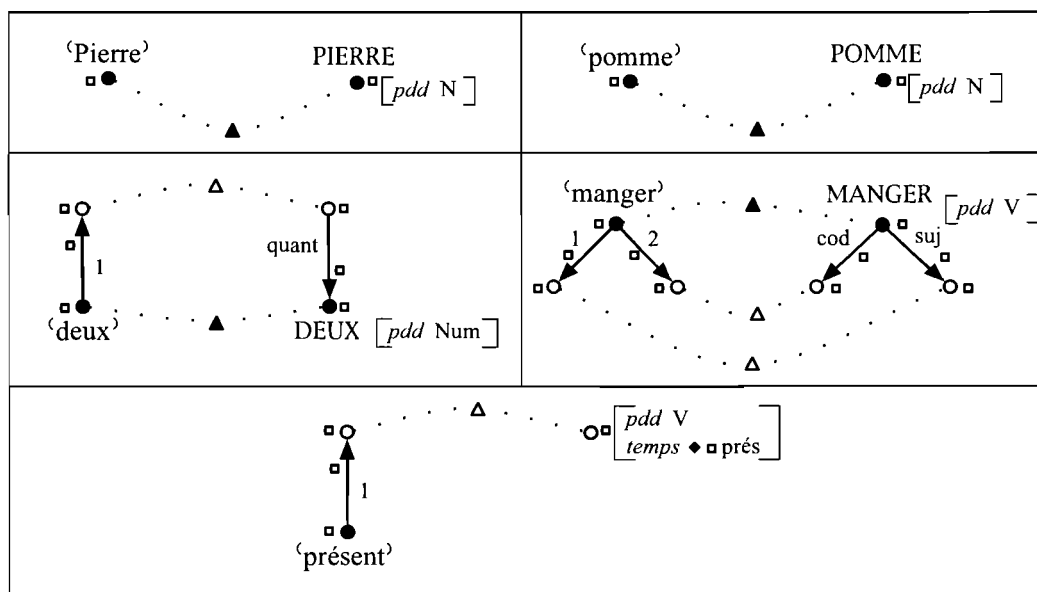


Figure 55 — Une interface sémantique-syntaxe articulée

Enfin, G_{synt} , avec p_{synt} comme polarité principale. La polarité $p_{sém-synt}$ est indiquée à la gauche des objets, tandis que la polarité p_{arbre} est masquée. Remarquons que les grammèmes d'accord ont une polarité $p_{sém-synt}$ saturée (il s'agit d'une solution temporaire; nous en verrons une autre au chapitre suivant, Ch. VIII, § 2, p. 298 et Ch. VIII, § 3, p. 316).

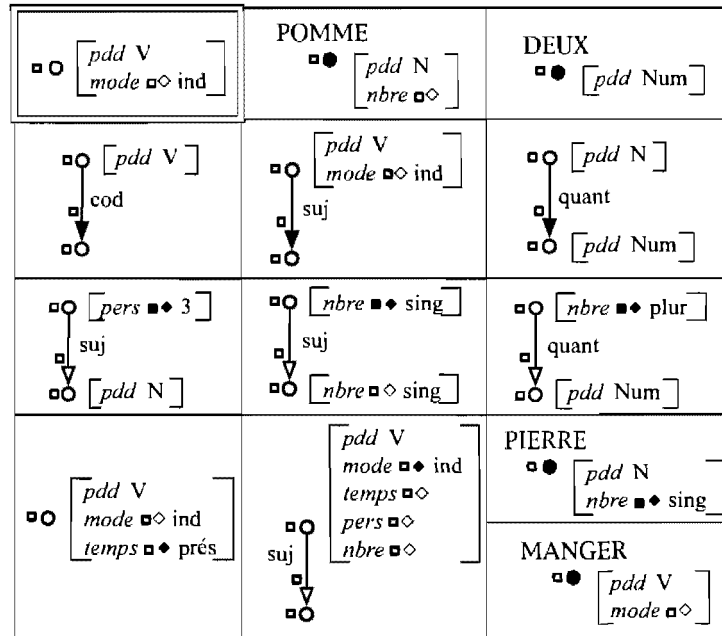


Figure 56 — Une grammaire syntaxique articulée

Ces trois petites grammaires s'appellent les unes les autres grâce au jeu des polarités d'articulation et peuvent produire, par exemple, la structure suivante.

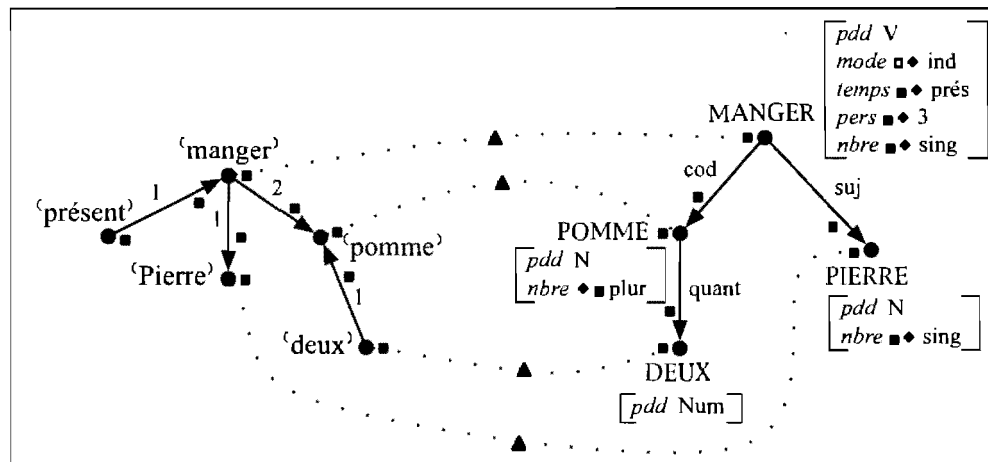


Figure 57 — Une structure de correspondance

Nous attirons l'attention du lecteur sur le grammème de mode que porte le lexème MANGER dans cette structure. C'est le seul qui ne soit pas saturé en polarité $p_{\text{sém-synt}}$. Cela est dû au fait qu'il nous manque la règle de $I_{\text{sém-synt}}$ qui décrit ce grammème superficiel. Nous avons vu que l'indicatif est vide de sens (cf. Ch. VI,

§ 3.1, p. 135). Nous verrons au prochain chapitre comment cela se modélise dans $\mathcal{I}_{\text{sém-synt}}$ (cf. Ch. VIII, § 3.2.2, p. 321).

Jusqu'ici nous n'avons considéré que le mode génératif de l'interface sémantique-syntaxe. Ce qui nous intéresse cependant, c'est le fonctionnement transductif de la grammaire, puisque c'est le seul qui puisse vraiment modéliser les processus linguistiques de synthèse et d'analyse. Considérons la synthèse. On donne en entrée une représentation sémantique et on cherche à obtenir toutes les représentations syntaxiques bien formées qui lui correspondent. Afin de déclencher l'application des règles de la grammaire, il s'agit simplement d'assigner une polarité $p_{\text{sém}}$ blanche à tous les objets de la représentation sémantique donnée en entrée. Par exemple, supposons que nous avons comme structure de départ²⁷ la représentation sémantique suivante :

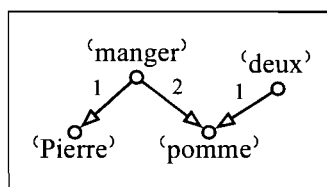


Figure 58 — Une structure sémantique à neutraliser

La neutralisation de ces polarités blanche ne pourra se faire qu'en combinant la structure sémantique à différentes règles de $\mathcal{G}_{\text{sém}}$ (qui elles contiennent des objets de polarité $p_{\text{sém}}$ noire). Ce faisant, on en vérifie la bonne formation et on obtient la structure suivante :

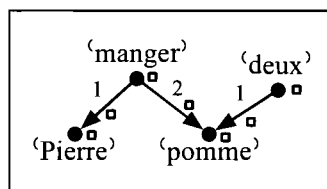


Figure 59 — Une structure sémantique à mettre en correspondance

27. On prendra soin de ne pas utiliser l'expression «structure initiale» pour parler de la structure donnée en entrée pour éviter la confusion avec les structures initiales des grammaires (cf. § 2, p. 226).

Puisque $G_{sém}$ est articulée avec $I_{sém-synt}$, l'application de ses règles a introduit de nouvelles polarités $p_{sém-synt}$ blanches, comme on peut le voir ci-dessus. Elles doivent à leur tour être neutralisées par les règles de $I_{sém-synt}$. Ces règles construisent les objets syntaxiques correspondant aux objets sémantiques et leur assignent une polarité p_{synt} blanche. On obtient ainsi la structure suivante :

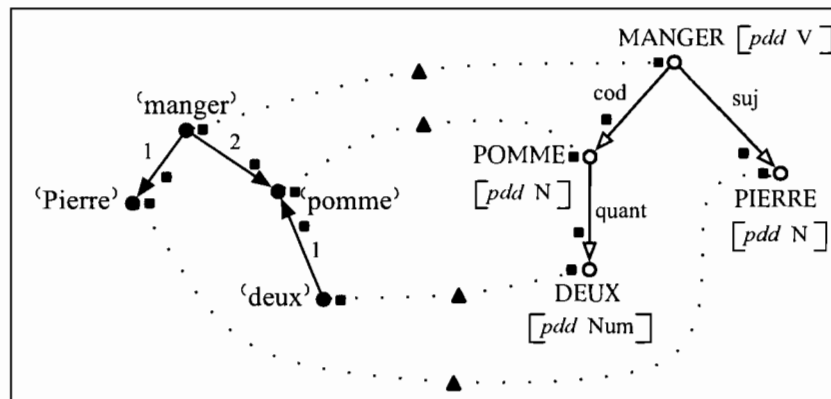


Figure 60 — Une structure syntaxique à valider

Ces polarités blanches demandent à être neutralisées par G_{synt} . L'application de G_{synt} déclenchera celle de G_{arbre} , puisque ces grammaires sont articulées. Cette dernière pourra vérifier que la structure syntaxique obtenue est vraiment un arbre. Toutefois, du point de vue de G_{synt} , elle n'est pas tout à fait bien formée. En effet, G_{synt} ne peut pas construire exactement cette structure puisque la règle qui construit le nœud MANGER exige la présence d'un mode pour ce lexème. Supposons, pour simplifier le problème, que les modes n'expriment aucun sens. On pourra donc choisir par exemple d'ajouter le grammème de l'indicatif, puisque notre petite grammaire syntaxique comporte une règle qui dit que l'indicatif est un mode valide pour un verbe. Cependant, la même règle introduit également des grammèmes de personne, de nombre et de temps polarisés en blanc. Les règles d'accord suffisent à saturer les grammèmes de personne et de nombre, mais le grammème de temps ne peut pas être neutralisé seulement par les règles de G_{synt} (celui de mode non plus, mais nous laisserons cette question de côté). On a alors, après l'application de G_{synt} , la structure non neutre de la Figure 61 ci-dessous où les grammèmes de mode et de temps n'ont pas de correspondant au niveau sémantique.

VII. La grammaire d'unification Sens-Texte

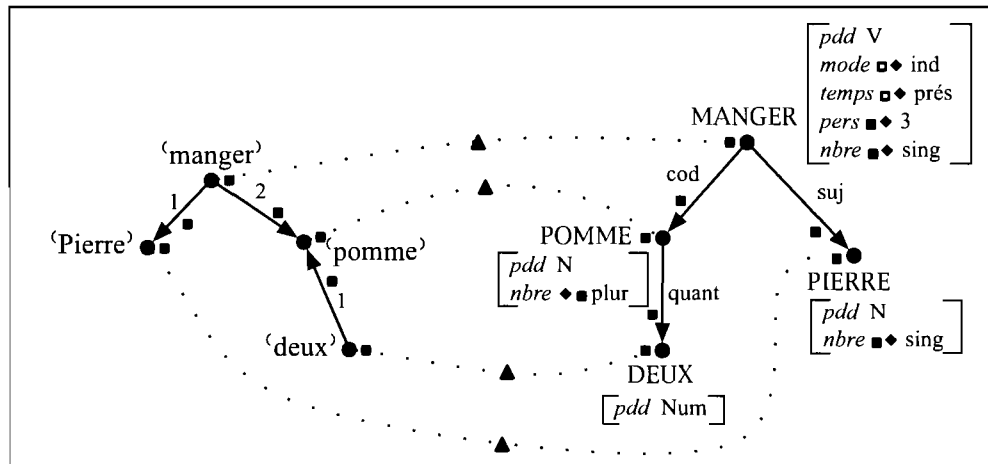


Figure 61 — Deux grammèmes à mettre en correspondance

Seules les règles de $I_{\text{sém-synt}}$ pourraient neutraliser ces polarités. Mais cela ne pose pas vraiment de problème puisque les règles ne sont pas orientées et les modules ne s'appliquent pas forcément dans un ordre séquentiel. Les grammaires GUST présentent en effet une caractéristique particulièrement intéressante qui les distinguent des grammaires de correspondance habituelles. En général, une grammaire transductive ne fait que construire une structure correspondant à celle donnée en entrée, sans jamais modifier cette dernière. Or, les grammaires GUST peuvent compléter la structure donnée en entrée si elle ne peut pas être mise telle quelle en correspondance avec une structure bien formée. On pourra donc neutraliser le grammème de temps, par exemple, en appliquant dans l'autre direction la règle de correspondance entre le sémantème 'présent' et le grammème qui l'exprime (introduite à la Figure 55, p. 261) ainsi que la règle sémantique qui construit le sémantème 'présent' et le relie à son actant (introduite à la Figure 54, p. 261). L'application de ces deux règles nous donne exactement la structure de correspondance de la Figure 57 (p. 262), où apparaît un sémantème 'présent' qui n'était pas dans la représentation sémantique à l'origine. Il ne reste alors plus qu'à neutraliser le grammème de mode en appliquant des règles équivalentes, ce qui ajoutera encore un élément de sens dans la représentation sémantique.

Ainsi, le jeu des polarités de GUST permet de modifier en cours de route la représentation donnée en entrée afin de satisfaire des contraintes provenant

d'autres niveaux de représentation. Bien sûr, il peut y avoir différentes façons de satisfaire ces contraintes et on obtiendra autant de structures sémantiques modifiées qu'il y a de combinaisons de règles qui permettent de neutraliser toutes les polarités. Dans notre exemple, nous avons choisi le temps présent, mais nous aurions pu en choisir un autre (en supposant que notre grammaire contienne des règles équivalentes pour les autres temps). Dans le cadre d'un système de génération automatique de texte, le programme devrait alors choisir parmi les graphes sémantiques proposés celui qui convient le mieux selon la situation. En d'autres termes, il devrait répondre à la question «Quand Pierre mange-t-il deux pommes: est-ce dans le passé, dans le présent ou dans le futur?».

Il faut noter que les contraintes décrites dans les grammaires de bonne formation peuvent avoir des répercussions non seulement sur les niveaux de représentation adjacents, comme on vient de le voir, mais également sur des niveaux plus éloignés. Par exemple, la grammaire de bonne formation phonologique doit contenir des règles qui force la superposition d'une prosodie à toute chaîne de phonèmes. S'il est vrai que cette prosodie peut être calculée au moins en partie d'après la représentation syntaxique ou morphotopologique, certains de ses éléments sont néanmoins porteurs de sens et doivent avoir un correspondant au niveau sémantique. La satisfaction de contraintes phonologiques peut donc entraîner la modification des structures plus profondes, et ce, jusqu'au niveau sémantique.

Ce «va-et-vient grammatical» est une des caractéristiques les plus intéressantes de ce formalisme puisqu'elle permet de modéliser de façon très élégante le fait que les sens grammaticaux ne sont pas nécessairement exprimés en réponse à un besoin communicatif de la part du locuteur, mais plutôt parce qu'ils sont imposés par la langue (voir le Chapitre IV à ce sujet). Ce type de mécanisme, en plus d'être plausible d'un point de vue cognitif, pourrait s'avérer particulièrement utile dans le cadre de la traduction automatique, où la représentation sémantique obtenue par analyse dans la langue source n'est pas forcément exprimable telle quelle dans la langue cible. En principe, le même mécanisme devrait également opérer en analyse. Lorsque la chaîne sonore donnée en entrée ne peut pas aboutir à une représen-

tation sémantique bien formée, la grammaire devrait fournir des chaînes sonore corrigées, exactement comme un locuteur peut essayer de «deviner» ce qui a été dit quand il n'a pas bien entendu. Cependant, comme nous n'avons pas encore une idée précise de ce à quoi ressembleront les représentations phonologiques en GUST, il n'est pas possible pour l'instant de vérifier cette hypothèse.

Pour conclure, on a vu que la neutralisation des polarités $p_{sém}$ blanches de la structure de départ déclenchait une chaîne de neutralisations qui ont construit tous les objets jusqu'au niveau de représentation le plus éloigné grâce au mécanisme d'articulation des modules, et ce tout en corrigeant la structure de départ qui s'était avérée incomplète. Le modèle que nous présentons ici ne va pas plus loin que le niveau syntaxique, mais on peut ajouter autant de modules qu'on le souhaite en utilisant toujours la même stratégie d'articulation entre chacun.

Ajoutons en terminant que nous n'avons considéré dans ce chapitre qu'une architecture en pipe-line, héritée de la TST. Mais en fait, rien n'empêche, ni dans le formalisme, ni dans la théorie, de mettre en correspondance des éléments de niveaux qui ne sont pas adjacents. Par exemple, un sémantème qui s'exprime par une intonation peut (et doit) être mis directement en correspondance avec le prosodème qui l'exprime. Cependant, comme nous ne nous intéressons dans cette thèse qu'à l'interface sémantique-syntaxe, nous laissons de côté cette discussion.

6 La décomposition sémantique en GUST

La modélisation d'une langue donnée ne saurait être complète sans la description du sens de ses signes sémantiquement pleins²⁸, qu'ils soient lexicaux ou grammaticaux. Dans le cadre de la TST, cela se fait par des règles d'équivalence entre des configurations de sémantèmes. Par exemple, la règle ci-dessous donne la décomposition du sémantème 'Torontoise'.

28. Le signifié d'un signe linguistique n'est pas forcément d'ordre sémantique (voir à ce sujet Mel'čuk 1993). Par exemple, le signifié d'un grammème d'accord est la relation syntaxique qu'il signale entre deux lexèmes. Nous ne considérons ici que les signes dont le signifié est au moins partiellement descriptible par une configuration de sémantèmes.

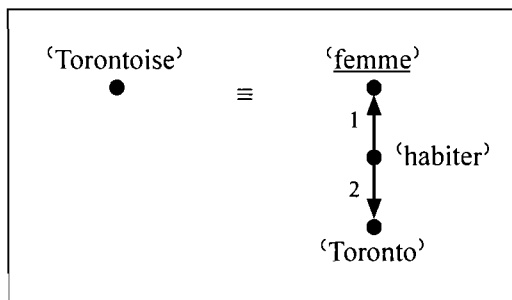


Figure 62 — Une règle d'équivalence sémantique de la TST

Ces règles sont nécessaires pour modéliser le paraphrasage, un phénomène central dans la langue. Par exemple, elles permettent de rendre compte de l'équivalence sémantique qui existe entre les trois groupes nominaux suivants :

- (152) a. *une Torontoise rousse*
 b. *une rousse qui habite Toronto*
 c. *une femme rousse qui habite Toronto*

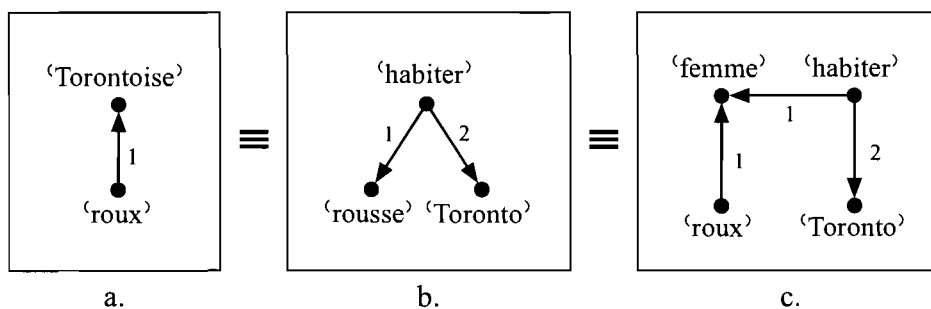


Figure 63 — Trois paraphrases qui n'ont pas la même RSém

Ces trois énoncés expriment exactement les mêmes sémantèmes, soit ceux visibles à la Figure 63c, mais ils n'encodent pas exactement le même message puisqu'ils diffèrent par la façon dont ils regroupent ces sémantèmes. Dans le premier énoncé, les sens 'femme', 'habiter' et 'Toronto' sont exprimés par un même nom, TORONTOISE (*cf.* Figure 63a), alors que dans le second ce sont 'femme' et 'roux' qui sont regroupés et exprimés par le nom ROUSSE (*cf.* Figure 63b), tandis que dans le troisième, à chaque sémantème correspond un lexème distinct (*cf.* Figure 63c).

Si ces trois énoncés n'ont pas la même représentation sémantique, c'est en vertu du principe du bloc maximal (*cf.* Mel'čuk 1988b, 1989, 1993; Polguère 1990), qui

veut que la représentation sémantique canonique d'un énoncé soit réduite au maximum. Ce principe vise à éviter l'arbitraire; en effet, sans ce principe, on ne saurait où arrêter la décomposition dans la représentation sémantique d'un énoncé²⁹.

L'équivalence entre les trois représentations sémantiques de la Figure 63 est assurée par des règles d'expansion / réduction comme celle de la Figure 62³⁰. Ces règles sont également nécessaires si on souhaite avoir accès à la structure interne d'un sémantème. Notamment, on peut vouloir vérifier que les actants d'un prédicat donné soient d'un certain type sémantique (par exemple, que le premier actant de 'manger' dénote un être animé, *cf.* § 5.1.1, p. 239). C'est possible de le faire avec des règles d'équivalence. Toutefois, ces règles soulèvent au moins trois questions quant à leur formalisation:

- 1) Où se situent les règles d'équivalence sémantique par rapport aux différents modules d'un modèle linguistique?
- 2) Comment peuvent-elles interagir avec les autres règles du modèle?
- 3) Peuvent-elles être formalisées en GUST, et si oui, comment?

Nous tentons d'y répondre dans cette section, qui reprend l'article de Lareau (2007), avec quelques modifications. Nous ne cherchons pas à déterminer précisément le rôle de la décomposition sémantique. Nous passons également sous silence le problème complexe de la sélection lexicale³¹, pourtant étroitement lié à la décomposition sémantique. Nous nous situons au niveau de la structure des règles de décomposition et de leur place dans un modèle déclaratif global de la langue.

Dans le cadre de la TST, les règles d'équivalence sémantique ne font pas partie des modules de transition qui constituent un modèle Sens-Texte, puisqu'elles encodent des phénomènes orthogonaux aux processus de synthèse et d'analyse simulés

29. La seule autre solution possible, si on veut éviter l'arbitraire, est de toujours décomposer jusqu'aux sens primitifs. C'est la position défendue notamment par Wierzbicka (1996), mais elle pose des problèmes théoriques (comment établir la liste des primitifs?) et pratiques (les structures deviennent illisibles). Voir à ce propos Mel'čuk (1989).

30. Pour une discussion de ce type de règle, voir notamment Polguère (1990), Kahane & Mel'čuk (1999) et Milićević (2007).

31. Voir à ce propos Polguère (1990, 1998b) et Wanner (1992, 1997).

par les règles de correspondance; on les trouve dans le module de paraphrasage sémantique. Pour les mêmes raisons, ces règles ne peuvent pas constituer une grammaire d'interface en GUST, puisqu'elles ne mettent pas en relation des objets de niveaux adjacents. Plutôt que de regrouper ces règles dans un module distinct orthogonal aux processus de synthèse et d'analyse, nous aimerions les intégrer à la grammaire de bonne formation sémantique de GUST.

Afin d'intégrer les règles d'expansion / réduction, nous devons modifier notre première caractérisation de la grammaire de bonne formation sémantique, $\mathcal{G}_{\text{sém}}$ (cf. § 5.1.1, p. 239). Les nœuds sémantiques pourront dorénavant être associés par une fonction à un ensemble de nœuds eux-même reliés par des arcs qui forment de petites structures sémantiques représentant les constituants du sémantème ainsi décomposé. Les nœuds qui décomposent un autre nœud sont reliés à ce dernier par la fonction *réduction*. Parmi les nœuds qui forment la décomposition d'un sémantème, un seul est identifié comme le sens principal. Il s'agit du genre prochain et il est identifié par la fonction *générique*, qui prend comme argument un sémantème (un nœud) et retourne son genre prochain (un autre nœud)³². Graphiquement, nous représentons les décompositions à l'intérieur d'une bulle en pointillés liée au nœud réduit, dans laquelle le genre prochain est souligné. Par souci de lisibilité, nous réduisons la taille des caractères et des traits à l'intérieur de ces bulles.

La figure suivante encode la règle d'équivalence 'Torontoise' = 'femme qui habite Toronto', tant sous forme explicite que sous forme graphique. Toutes les figures qui suivent dans la présente section montrent explicitement les deux polarités de $\mathcal{G}_{\text{sém}}$ pour chaque objet: $p_{\text{sém}}$ (qui contrôle la saturation des structures sémantiques) est indiquée par la couleur de l'objet alors que $p_{\text{sém-synt}}$ (qui assure l'interface avec $\mathcal{I}_{\text{sém-synt}}$) est indiquée par un petit carré à côté de l'objet. Par ailleurs, dans la Figure 64, nous ajoutons aux nœuds de la version graphique, entre parenthèses et en italique, les numéros d'identification utilisés dans la version tex-

32. Il n'est pas clair que l'on puisse toujours identifier un seul sémantème comme genre prochain pour tous les sens de la langue. Par exemple, Polguère (2003b) mentionne le cas de 'fruit', qui peut être défini à la fois par 'nourriture' et 'partie de plante'. Ce problème dépasse le cadre de notre thèse.

tuelle afin que le lecteur puisse plus facilement mettre en relation les deux formes de représentation.

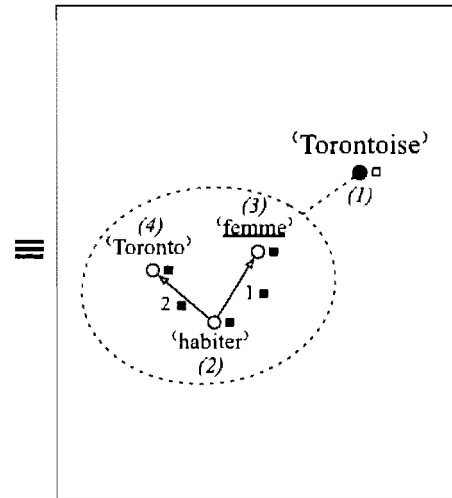
$$\langle \{ 1:\text{nœud}_{\text{sém}}, 2:\text{nœud}_{\text{sém}}, 3:\text{nœud}_{\text{sém}}, \\ 4:\text{nœud}_{\text{sém}}, 5:\text{arc}_{\text{sém}}, 6:\text{arc}_{\text{sém}} \}, \\ \{ \text{source}(5)=2, \text{cible}(5)=3, \text{source}(6)=2, \\ \text{cible}(6)=4, \text{étiquette}(1)=\text{"Torontoise"}, \\ \text{étiquette}(2)=\text{"habiter"}, \text{étiquette}(3)=\text{"femme"}, \\ \text{étiquette}(4)=\text{"Toronto"}, \text{étiquette}(5)=\text{"1"}, \\ \text{étiquette}(6)=\text{"2"} \}, \\ p_{\text{sém}}(1)=\blacksquare, p_{\text{sém}}(2)=\square, p_{\text{sém}}(3)=\square, \\ p_{\text{sém}}(4)=\square, p_{\text{sém}}(5)=\square, p_{\text{sém}}(6)=\square, \\ p_{\text{sém-synt}}(1)=\square, p_{\text{sém-synt}}(2)=\blacksquare, p_{\text{sém-synt}}(3)=\blacksquare, \\ p_{\text{sém-synt}}(4)=\blacksquare, p_{\text{sém-synt}}(5)=\blacksquare, p_{\text{sém-synt}}(6)=\blacksquare \\ \text{générique}(1)=2, \text{réduction}(2)=1, \\ \text{réduction}(3)=1, \text{réduction}(4)=1 \} \rangle$$


Figure 64 — La décomposition sémantique de 'Torontoise'

Les sémantèmes 'rousse' et 'femme' peuvent se décomposer de la même façon, comme le montre la Figure 65. On remarquera que les nœuds à l'intérieur des décompositions portent toujours une polarité $p_{\text{sém}}$ blanche. Ils forcent ainsi l'application d'autres règles du même module pour vérifier leur bonne formation. Par exemple, le nœud 'femme' apparaissant dans la décomposition de 'rousse' forcera l'application de la règle qui en vérifie la bonne formation (à droite dans la même figure), introduisant par le fait même la décomposition de ce sémantème.

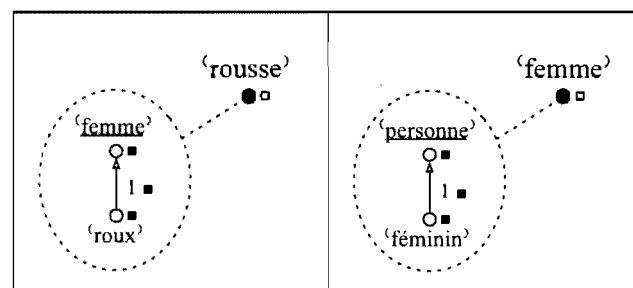


Figure 65 — La décomposition sémantique de 'rousse' et de 'femme'

On remarquera également que les nœuds à l'intérieur des décompositions sont toujours saturés en $p_{\text{sém-synt}}$. Cela a pour but d'éviter de réaliser à la fois un sens et sa définition dans un même énoncé. Lorsqu'un fragment de la structure sémantique

a été réduit, tous ses éléments deviennent invisibles pour l'interface $I_{\text{sém-synt}}$, qui ne peut plus voir que le sémantème réduit.

Désormais, toutes les règles de $G_{\text{sém}}$ qui construisent un nœud devront également en fournir la décomposition. Seuls les primitifs sémantiques font exception à cette règle, puisqu'ils ne peuvent pas être décomposés. Dans un système concret de traitement automatique de la langue, il n'est certainement pas souhaitable de toujours décomposer les sémantèmes jusqu'aux primitifs, puisque cela entraînerait beaucoup de calcul inutile. On pourra alors avoir, pour chaque sémantème décomposable, deux règles : une qui le traite comme un primitif, et une autre qui en donne la décomposition. Pour les besoins de notre exposé, nous allons considérer que '(de sexe) féminin', '(aux cheveux) roux', 'habiter', 'Toronto' et 'personne' sont des sens indécomposables, ce qui n'est évidemment pas le cas en réalité.

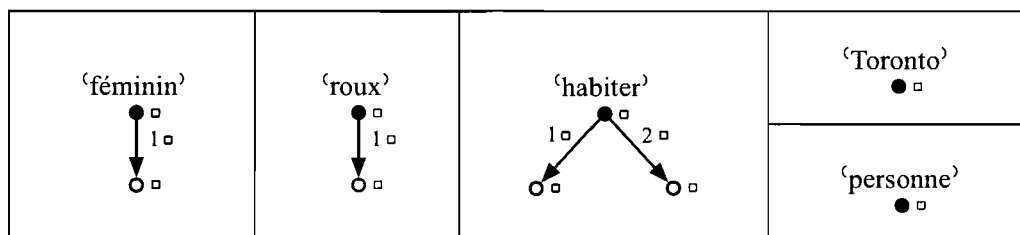


Figure 66 — Les sens «primitifs» dans la grammaire sémantique

Il est intéressant de noter que nos règles de décomposition ne forment pas une grammaire de correspondance, contrairement au module de paraphrasage sémantique de la TST. En effet, nos règles de décomposition ne mettent pas en correspondance une représentation sémantique et les représentations décomposées équivalentes. Il ne s'agit pourtant pas d'une grammaire transformationnelle non plus, puisque la structure de départ n'est pas modifiée.

Voyons maintenant comment nos règles peuvent s'appliquer en synthèse. Le processus de synthèse peut être lancé à partir de n'importe quelle représentation, réduite ou non. Dans notre exemple, nous partirons d'une représentation décomposée «au maximum». D'abord, les sens décrits à la Figure 66, considérés ici comme «primitifs» (bien que ce ne soit pas réellement le cas) nous permettent de construire la représentation de la Figure 67 ci-dessous.

VII. La grammaire d'unification Sens-Texte

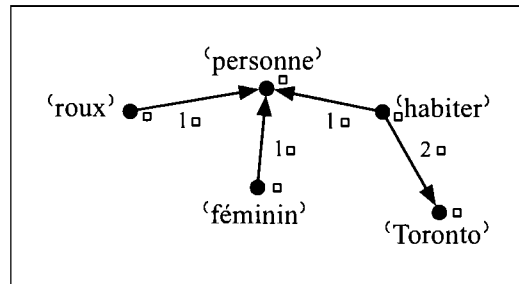


Figure 67 — Une représentation sémantique décomposée à réaliser

Il est parfaitement possible d'utiliser cette représentation sémantique telle quelle et de la mettre en correspondance avec un arbre syntaxique sans la réduire aucunement. On obtient alors l'énoncé (153):

(153) *Une personne rousse de sexe féminin qui habite Toronto.*

On peut également opérer une première réduction en appliquant la règle qui donne la définition de 'femme' (cf. Figure 65), ce qui nous donne la représentation suivante:

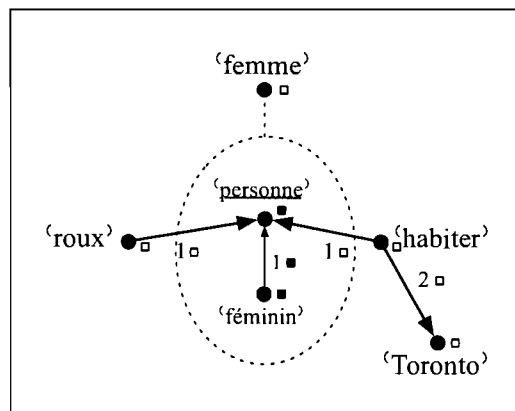


Figure 68 — La réduction de 'personne (de sexe) féminin'

Cette représentation ne peut toutefois pas être mise en correspondance avec un arbre syntaxique puisque du point de vue de l'interface sémantique-syntaxe elle n'est pas connexe ('personne' étant saturé en $p_{\text{sém-synt}}$, il n'est plus visible pour l'interface). La structure qui en résulterait ne serait pas un arbre et ne pourrait donc pas être saturée en p_{arbre} (cf. § 5.1.2, p. 242). Pour rendre la structure connexe, il faut en quelque sorte «déplacer» les relations sémantiques qui pointaient vers 'personne', puisque ce sémantème n'est plus visible pour l'interface sémantique-

syntaxe. Les prédicats qui portaient sur ce sémantème doivent maintenant porter sur ‘femme’. Kahane & Mel’čuk (1999) proposaient une métarègle qui recopie les relations sémantiques pointant vers un nœud pour qu’elles pointent vers le sémantème réduit dont ce nœud est le genre prochain³³. Cette règle se formalise en GUST de la façon suivante :

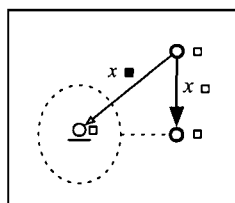


Figure 69 — Une règle de réancrage du gouverneur

Une fois cette règle appliquée, on obtient la structure suivante :

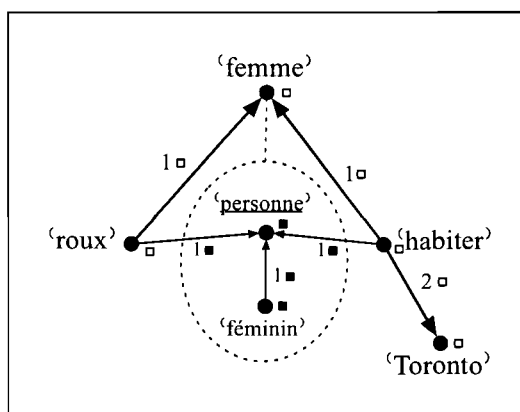


Figure 70 — La réduction de ‘personne (de sexe) féminin’ avec réancrage des gouverneurs

Les relations pointant vers ‘personne’ ont été « neutralisées » du point de vue de l’interface sémantique-syntaxe par la saturation de leur polarité $p_{\text{sém-synt}}$. La structure résultante peut tout de suite être mise en correspondance avec un arbre syntaxique, ce qui donne l’énoncé (154) :

(154) *Une femme rousse qui habite Toronto.*

33. Cette règle ne concerne que le genre prochain. En effet, si une relation pointe vers un sémantème qui n’est pas le genre prochain, alors la réduction n’est pas possible (sauf, comme le remarquait Mel’čuk 2004, pour les expressions phraséologisées [*payer les yeux de la tête*], mais elles doivent être décrites par des règles spécifiques). Par exemple, ‘femme très rousse’ ne peut pas être réduit sans perte d’information [**une très rousse_(N)*], alors que ‘grande femme rousse’ ne pose aucun problème [*une grande rousse_(N)*].

Elle peut aussi être réduite encore une fois en appliquant la règle qui construit 'rousse'. Cela nous donne la structure en 71a, puis celle en 71b une fois la règle de réancrage des gouverneurs appliquée.

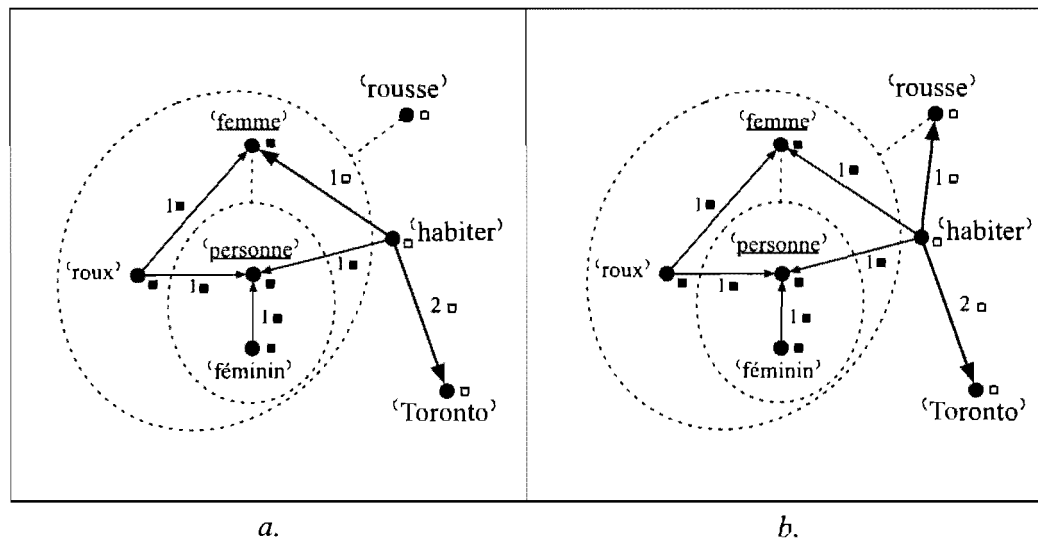


Figure 71 — La réduction de 'femme rousse' avec réancrage des gouverneurs

Il résulte que les seuls sémantèmes accessibles à $I_{\text{sém-synt}}$ (donc demandant à être exprimés), sont 'rousse', 'habiter' et 'Toronto', ce qui permet de former l'énoncé (4).

(155) *Une rousse qui habite Toronto.*

Bien entendu, ce n'est qu'une des réductions possibles. On peut réduire ces configurations de sémantèmes autrement en utilisant d'autres règles. Dans tous les cas, on retrouvera toujours la configuration sémantique de départ (cf. Figure 67, p. 273), mais de nouveaux sémantèmes réduits apparaîtront et les sens accessibles à $I_{\text{sém-synt}}$ ne seront pas les mêmes.

Ainsi, le type de règle que nous proposons permet, sans aucune procédure particulière et sans avoir à ajouter ni nouveau module ni nouvelle polarité au modèle, de trouver automatiquement toutes les réductions possibles pour une représentation sémantique donnée, et donc toutes ses différentes expressions. Nous avons illustré le fonctionnement de ces règles en synthèse, mais elles fonctionnent également en analyse. Elles fournissent alors, pour la phrase analysée, non seulement sa

représentation sémantique réduite, mais également sa décomposition complète ainsi que toutes les décompositions intermédiaires. Cette caractéristique du formalisme proposé peut s'avérer fort intéressante notamment en traduction automatique, puisque la traduction des sens lexicaux et grammaticaux doit souvent être approximative; il faut donc avoir accès à leur décomposition.

En outre, le formalisme que nous suggérons ici permet de représenter de façon assez naturelle le typage des actants sémantiques (*cf.* § 5.1.1, p. 239). Si on souhaite restreindre la combinatoire sémantique d'un prédicat, on peut en étiqueter les actants dans la règle de $G_{\text{sém}}$ qui le construit. Par exemple, si on veut que le prédicat 'habiter' ne puisse prendre comme premier actant qu'un sémantème dénotant une personne et comme deuxième actant un sémantème qui dénote un lieu, il suffit de modifier la règle (que nous avons déjà vue à la Figure 66, p. 272) qui introduit ce sémantème de la façon suivante :

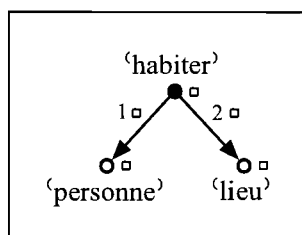


Figure 72 — Le typage sémantique des actants en GUST

Sans avoir accès à la décomposition du sens, le premier actant de 'habiter' ne pourrait pas, par exemple, être 'rousse' puisque l'étiquette 'personne' du nœud en Figure 72 ne peut pas directement s'unifier avec l'étiquette 'rousse'. Par contre, avec le type de représentation que nous venons de présenter, puisque 'personne' est le genre prochain de 'femme', qui lui-même est le genre prochain de 'rousse', ce dernier pourra, grâce à la règle de réancrage des gouverneurs, être le premier actant de 'habiter', sans qu'il ne soit nécessaire d'activer une quelconque procédure ou de faire appel à une structure de donnée séparée qui encode une hiérarchie des types sémantiques. Comme on peut l'observer à la Figure 71b, le formalisme permet en effet de vérifier que 'personne' est une composante centrale du premier actant de

‘habiter’, et ce peu importe à quelle «profondeur» se trouve enfouie cette composante dans le sens de l’actant³⁴.

Insistons sur le fait que le formalisme n’oblige pas à typer les actants d’un sémantème. On peut très bien ne pas étiqueter les actants dans les règles qui introduisent les sémantèmes prédicatifs sans que n’en souffre la grammaire. Le modèle permet alors de mettre en correspondance des phrases absurdes avec des représentations sémantiques formellement correctes, mais difficiles à interpréter. Revient-il à la grammaire de rejeter de telles phrases? D’un côté, il est vrai que la langue permet d’exprimer même des idées absurdes ou contradictoires. Si la grammaire n’est qu’un modèle fonctionnel de la langue, alors elle doit permettre de telles phrases. D’un autre côté, le typage des actants semble modéliser une certaine connaissance que les locuteurs ont des sens de la langue. De fait, dans le cadre de la TST, Mel’čuk et Polguère n’hésitent pas à y avoir recours dans leur base de données lexicales *DiCo* (Polguère 2003a). En outre, cela peut s’avérer très utile en linguistique computationnelle, par exemple pour valider une analyse incertaine. Enfin, la question dépasse le cadre de notre thèse, mais quoi qu’il en soit, le formalisme permet d’implémenter, au moins dans une certaine mesure, les deux solutions: le linguiste choisira à sa guise de typer ou non les actants des prédicats qu’il décrit dans la grammaire de bonne formation sémantique.

Nous avons illustré le fonctionnement des règles de décomposition sémantique à l’aide de sémantèmes lexicaux. Il va de soi que ce sont les sens grammaticaux qui nous intéressent dans cette thèse. Nous allons maintenant voir au Chapitre VIII que le même mécanisme peut être utilisé pour décrire le sens des grammèmes verbaux en français.

34. Bien entendu, puisque nous n’avons pas fourni dans nos règles la décomposition sémantique de ‘Toronto’, le mécanisme ne peut pas fonctionner pour le deuxième actant de ‘habiter’, qui ne pourra pas s’unifier à ‘Toronto’. Mais si nous complétons notre grammaire pour fournir la décomposition de ‘Toronto’ et encore de ses composantes sémantiques, ‘habiter’ pourra alors l’accepter comme deuxième actant puisque ‘Toronto’ \supset ‘ville’ \supset ‘lieu’.

7 Synthèse

GUST est une grammaire, à la fois générative, transductive et équative, qui établit la correspondance entre des sens et des formes. Alors que la TST décrit cette correspondance au moyen de sept niveaux de représentation, GUST n'en possède que quatre, puisque les niveaux profonds et de surface ne sont pas distingués. À chaque niveau de représentation (sémantique, syntaxique, morphotopologique et phonologique), une grammaire de bonne formation assure la construction de représentations valides. La correspondance entre les niveaux de représentation est assurée par trois grammaires d'interface. Les règles d'interface décrivent des signes (au sens large) dont le signifié et le signifiant sont des éléments des représentations mises en correspondances.

Toutes les règles de GUST sont des structures élémentaires qui, en se combinant, construisent les structures complexes qui représentent les différents niveaux d'organisation des énoncés. Afin d'assurer la construction de représentations complètes, les structures sont polarisées. En se combinant, les polarités se neutralisent; une structure est saturée quand tous ses objets ont été neutralisés.

Un même objet peut porter plus d'une polarité, ce qui permet d'articuler les modules. Ce mécanisme est particulièrement intéressant puisqu'il permet d'appliquer les règles dans n'importe quel ordre, et même de compléter une structure incomplète donnée en entrée.

Le Tableau XXVI ci-dessous récapitule les caractéristiques formelles de chaque module. Il indique les types d'objets que peuvent contenir chacun. À chaque type d'objet sont associés certaines fonctions, dont plusieurs sont obligatoires (elles sont indiquées en gras). Nous indiquons également le type de valeur retourné par chaque fonction.

Ce tableau montre explicitement quelles sont les réalités linguistiques représentées par les différents objets des structures GUST. Bien entendu, il ne faut pas confondre les signes linguistiques et leur représentation formelle. Mais nous nous

Module	Types d'objets	Fonctions	Phénomènes représentés
$G_{\text{sém}}$	$\text{nœud}_{\text{sém}}$	$p_{\text{sém}} \rightarrow \mathcal{P}$ $p_{\text{sém-synt}} \rightarrow \square$ <i>étiquette</i> $\rightarrow S$ <i>générique</i> $\rightarrow x : \text{nœud}_{\text{sém}}$ <i>réduction</i> $\rightarrow x : \text{nœud}_{\text{sém}}$	sémantèmes
	$\text{arc}_{\text{sém}}$	$p_{\text{sém}} \rightarrow \mathcal{P}$ $p_{\text{sém-synt}} \rightarrow \square$ <i>source</i> $\rightarrow x : \text{nœud}_{\text{sém}}$ <i>cible</i> $\rightarrow x : \text{nœud}_{\text{sém}}$ <i>étiquette</i> $\rightarrow \mathcal{R}_{\text{sém}}$	relations sémantiques
G_{synt}	$\text{nœud}_{\text{synt}}$	$p_{\text{synt}} \rightarrow \mathcal{P}$ $p_{\text{arbre}} \rightarrow \square$ $p_{\text{sém-synt}} \rightarrow \square$ $p_{\text{synt-topo}} \rightarrow \square$ <i>pdd</i> $\rightarrow C$ <i>étiquette</i> $\rightarrow L$ <i>genre</i> $\rightarrow \{\text{masc, fém}\}$ <i>groupe</i> $\rightarrow \{1, 2, 3, \text{irr.}\}$... (traits de syntactique) <i>mode</i> $\rightarrow x : \text{obj-gramm}$ <i>nombre</i> $\rightarrow x : \text{obj-gramm}$ <i>temps</i> $\rightarrow x : \text{obj-gramm}$... (catégories flexionnelles)	lexèmes
	arc_{synt}	$p_{\text{synt}} \rightarrow \mathcal{P}$ $p_{\text{arbre}} \rightarrow \square$ $p_{\text{sém-synt}} \rightarrow \square$ $p_{\text{synt-topo}} \rightarrow \square$ <i>source</i> $\rightarrow x : \text{nœud}_{\text{synt}}$ <i>cible</i> $\rightarrow x : \text{nœud}_{\text{synt}}$ <i>étiquette</i> $\rightarrow \mathcal{R}_{\text{synt}}$	relations syntaxiques
	obj-gramm	$p_{\text{synt}} \rightarrow \mathcal{P}$ $p_{\text{sém-synt}} \rightarrow \square$ $p_{\text{synt-topo}} \rightarrow \square$ <i>hôte</i> $\rightarrow x : \text{nœud}_{\text{synt}}$ <i>étiquette</i> $\rightarrow G$	grammèmes

Tableau XXVI — Tableau récapitulatif des modules de GUST et de leurs objets

G_{arbre}	$\text{nœud}_{\text{synt}}$	$p_{\text{arbre}} \rightarrow \mathcal{P}$ $p_{\text{synt}} \rightarrow \square$	lexèmes
	arc_{synt}	$p_{\text{arbre}} \rightarrow \mathcal{P}$ $p_{\text{synt}} \rightarrow \square$	relations syntaxiques
$I_{\text{sém-synt}}$	tous les types d'objets de $G_{\text{sém}}$ et G_{synt}	mêmes fonctions que dans $G_{\text{sém}}$ et G_{synt} sauf fonctions de polarité, plus: $p_{\text{sém-synt}} \rightarrow \mathcal{P}$ $p_{\text{sém}} \rightarrow \square$ (objets sémantiques) $p_{\text{synt}} \rightarrow \square$ (objets syntaxiques)	mêmes phénomènes
	correspondance	$p_{\text{sém-synt}} \rightarrow \mathcal{P}$ $\text{signifié} \rightarrow x: T_{\text{sém}}$ $\text{signifiant} \rightarrow x: T_{\text{synt}}/\text{obj-gramm}$	liens de correspondance

Tableau XXVI — Tableau récapitulatif des modules de GUST et de leurs objets

permettrons parfois quelques abus de langage afin d'alléger le texte. Par exemple, on trouvera dans cette thèse des expressions telles que «saturer un lexème» ou «cette règle introduit un grammème de finitude». Ces énoncés en soi sont incorrects, mais il faut les comprendre comme «saturer le nœud syntaxique qui représente un lexème» et «cette règle introduit un objet grammématique qui représente un grammème de finitude».

Finalement, nous proposons à l'Annexe B, à la fin de cette thèse, un encodage XML pour GUST. Ce type d'encodage est utile pour le traitement automatique des langues.

VIII. Un modèle formel de la conjugaison française

Nous avons déterminé au Chapitre VI les catégories flexionnelles verbales du français et les grammèmes qu'elles contiennent. Afin de pouvoir formaliser ce modèle, nous avons présenté et développé au Chapitre VII le formalisme GUST. Dans le présent chapitre, nous allons voir comment notre modèle de la flexion verbale française se représente dans ce formalisme.

Rappelons que ce sont essentiellement les grammèmes temporels qui nous intéressent dans cette thèse. Nous avons deux catégories flexionnelles temporelles, obligatoires pour tous les verbes à l'indicatif:

- 1) Les grammèmes de la catégorie de temps situent les faits par rapport à un point de repère. Il y a trois grammèmes dans cette catégorie: **simultané** (le fait est simultané à son point de repère), **antérieur** (le fait est antérieur à son point de repère) et **postérieur** (le fait est postérieur à son point de repère).
- 2) Les grammèmes de la catégorie de décalage situent un point de repère par rapport au moment d'énonciation. Il y a deux grammèmes de décalage: **non-décalé** (le repère est 'maintenant', c'est-à-dire que le fait est situé directement par rapport au moment d'énonciation, ou encore le repère est dans le futur) et **décalé** (le repère est situé dans le passé).

À ces deux catégories flexionnelles temporelles s'ajoute une classe de marqueurs de phase aspectuelle: AVOIR_{acc} (accompli), ¹EN TRAIN¹ (progressif), ALLER (prospectif), VENIR (passé récent) et ¹SUR LE POINT¹ (imminence), qui se combinent avec les grammèmes de temps et de décalage. Les grammèmes des autres catégories flexionnelles (finitude, voix, mode, personne, nombre et genre) ne servent pas à exprimer des sens temporels ou aspectuels¹.

1. Sauf **participe-é**, qui peut exprimer la résultativité.

Certains grammèmes ont plus d'une acception. Nous les avons décrites au Chapitre VI. Nous porterons une attention spéciale aux acceptions d'ordre temporel ou aspectuel; l'expression des sens modaux en français ne nous intéresse pas directement. Par ailleurs, nous ne prétendons pas à l'exhaustivité. Notre but n'est pas de faire l'inventaire de tous les sens que peuvent porter les grammèmes verbaux, mais simplement d'illustrer comment la polysémie grammaticale se représente en GUST, et comment notre modèle flexionnel permet de rendre compte des principaux sens de ces grammèmes.

Nous proposons des règles pertinentes pour la flexion verbale dans les modules de bonne formation sémantique ($G_{\text{sém}}$), de bonne formation syntaxique (G_{synt}) et d'interface sémantique-syntaxe ($I_{\text{sém-synt}}$) du français (Ch. VII, § 5, p. 236). Nous présenterons ces modules dans cet ordre, suivant la tradition au sein de la TST de privilégier la synthèse plutôt que l'analyse. Cette perspective ne signifie aucunement que notre modèle est orienté; il est tout à fait bi-directionnel.

Dans ce chapitre, toutes les règles sont numérotées, ce qui facilite les références. Pour les besoins de notre exposé, nous devons à l'occasion présenter des règles que nous ne souhaiterions pas conserver dans notre modèle. Afin d'éviter toute confusion, nous utilisons donc deux séries de numéros: les règles temporaires sont numérotées RT1, RT2, etc., alors que les règles qui font réellement partie de notre modèle linguistique sont marquées R1, R2, etc. Notons que les règles présentées au Chapitre VII ne l'étaient qu'à titre illustratif. Seules les règles du présent chapitre font partie de notre grammaire.

Sous chaque figure, nous indiquons en petits caractères le nom du module auquel appartiennent les règles qui y apparaissent. Nous donnons également le nom de la polarité principale. Le cas échéant, nous indiquons la ou les polarité(s) secondaire(s) entre parenthèses². Lorsque la valeur d'une polarité donnée pour tous les objets d'une règle est blanche, et seulement dans ce cas, cette polarité est

2. Rappelons que toutes les règles de GUST sont des structures multipolarisées et qu'elles peuvent donc toujours être présentées sous différents angles, selon la polarité qui est choisie comme polarité principale. Voir la discussion à ce propos p. 249.

occultée dans la figure. La plupart des règles ne montrent donc qu'une seule polarité, puisque la plupart n'enaturent qu'une seule à la fois.

1 La grammaire de bonne formation sémantique

Le rôle de la grammaire de bonne formation sémantique, $G_{\text{sém}}$, est de décrire les sens de la langue. Pour chaque sémantème, une règle de $G_{\text{sém}}$ indique s'il est un prédicat ou non et, le cas échéant, combien d'actants il doit avoir. C'est aussi dans ce module que sont données les définitions, par des décompositions sémantiques. L'encodage des décompositions sémantiques permet également de typer les actants d'un prédicat (*cf.* Ch. VII, § 6, p. 267).

Nous avons vu au Chapitre VI que le signifié des grammèmes verbaux du français n'est pas forcément représentable par une structure sémantique. En effet, les grammèmes peuvent exprimer des éléments d'autres structures :

- La structure communicative (pour les grammèmes profonds de voix).
- La structure référentielle (pour les valeurs d'irréel de l'imparfait et du subjonctif).
- La structure rhétorique (pour les acceptions atténuatives du futur, de l'imparfait et du conditionnel).

Nous n'allons pas discuter davantage ces signes, puisque nous n'avons pas étudié la formalisation de ces structures en GUST. De toute façon, ces signes ne concernent pas l'expression du temps en français; il ne sont donc pas d'un intérêt immédiat pour notre étude, et nous nous contenterons de ce que nous en avons dit dans les sections correspondantes du Chapitre VI. Par ailleurs, certaines grammies trouvent leur signifié dans la structure syntaxique. C'est le cas notamment des grammèmes d'accord et de la plupart des grammèmes de finitude et de mode. Ces signes ne sont pas pertinents pour $G_{\text{sém}}$.

Les autres sens exprimés par les grammèmes verbaux en français sont représentables, au moins partiellement, par une structure sémantique. Ils peuvent être

regroupés en trois familles: les sens temporels, les sens phasiques (ou aspectuels) et les sens modaux. C'est dans cet ordre que nous les aborderons, après avoir discuté brièvement du typage des actants sémantiques. Mais d'abord, nous ferons un bref rappel des caractéristiques formelles de $\mathcal{G}_{\text{sém}}$.

1.1 Les caractéristiques formelles de $\mathcal{G}_{\text{sém}}$

Nous rappelons ici les principales caractéristiques formelles de la grammaire de bonne formation sémantique. Pour une présentation plus détaillée, le lecteur pourra relire les sections correspondantes du chapitre précédent (Ch. VII, § 5.1.1, p. 239 et Ch. VII, § 6, p. 267).

Le rôle de $\mathcal{G}_{\text{sém}}$ est de vérifier la bonne formation des représentations sémantiques données en entrée (en synthèse) ou construites par la grammaire d'interface $\mathcal{I}_{\text{sém-synt}}$ (en analyse). Cela se fait par la saturation de la polarité $p_{\text{sém}}$.

Les objets de $\mathcal{G}_{\text{sém}}$ sont de deux types: nœuds sémantiques (qui représentent les sémantèmes) et arcs sémantiques (qui représentent les relations prédicat-argument entre les sémantèmes). Tous doivent porter une polarité $p_{\text{sém}}$, ainsi qu'une polarité d'articulation, $p_{\text{sém-synt}}$. Cette dernière est la plupart du temps blanche et nous ne l'indiquons dans les figures que lorsqu'elle est saturée.

Les règles de $\mathcal{G}_{\text{sém}}$ permettent également d'encoder les décompositions sémantiques. Cela se fait par le biais des fonctions *réduction* et *générique*. La première associe un sémantème à ses composantes sémantiques, alors que la seconde identifie parmi ces composantes le genre prochain. Les sémantèmes des décompositions sémantiques sont saturés en $p_{\text{sém-synt}}$, de façon à ce qu'ils ne soient pas réalisés en surface en même temps que le sémantème réduit dont ils sont une composante. Dans les figures, les décompositions sémantiques sont représentées à l'intérieur d'une bulle. Leurs arcs sont en pointillés et leurs nœuds sont plus petits, ce qui représente le fait que leur expression en syntaxe est bloquée.

1.2 Le typage des actants sémantiques

Les descriptions sémantiques que nous utilisons suivent l'usage en sémantique lexicale dans le cadre de la TST (*cf.* notamment Mel'čuk, Clas & Polguère 1995). En particulier, les actants sémantiques d'un prédicat sont typés (voir la discussion à ce propos, Ch. VII, § 5.1.1, p. 239). Le typage des actants modélise la combinatoire des sémantèmes en indiquant l'hyperonyme (direct ou indirect) des sémantèmes qui peuvent occuper une position donnée dans les places actantielles d'un prédicat. Cela se fait, dans le cadre de dictionnaires comme le *Dictionnaire explicatif et combinatoire du français contemporain* (Mel'čuk *et al.* 1984–1999), le *DiCo* (Polguère 2000) ou le *Lexique actif du français* (Mel'čuk & Polguère 2007) par exemple, par l'utilisation d'étiquettes sémantiques. Ces étiquettes sont des sens de la langue organisés dans une hiérarchie d'hyperonymes. Il s'agit donc d'une modélisation de l'organisation des *sens* linguistiques, et non de concepts plus ou moins universaux ou d'objets extra-linguistiques comme c'est le cas dans les ontologies et les taxinomies³.

Nous prenons comme cadre général pour nos descriptions sémantiques la hiérarchie des étiquettes du *Lexique actif du français*. Au sommet de cette hiérarchie, on trouve l'opposition entre faits et entités. Tous les verbes en français dénotent des faits⁴. Les faits et les entités se décomposent en divers types. La Figure 73, p. 286, montre à titre illustratif le lien entre 'quelque chose' et 'individu'. Cependant, pour les faits, nous ferons appel aux catégories de Vendler (1957) (états, activités, accomplissements et réalisations instantanées)⁵, puisqu'elles permettent de décrire de façon efficace la combinatoire des sémantèmes qui nous intéressent. Par contre, nous retiendrons la structuration de Mourelatos (1978), Figure 74, p. 286, puisqu'elle est basée sur des propriétés définitoires plutôt que sur des tests linguistiques comme c'était le cas dans la formulation initiale de Vendler⁶.

3. Voir Milićević (1997), Polguère (2003b) et Mel'čuk & Polguère (2007) pour une discussion détaillée.

4. C'est pourquoi nous avons utilisé partout dans cette thèse le terme « fait » plutôt que le terme plus courant de « procès » pour renvoyer à la dénotation des verbes.

5. *Cf.* Ch. VI, § 6.1.2, p. 181.

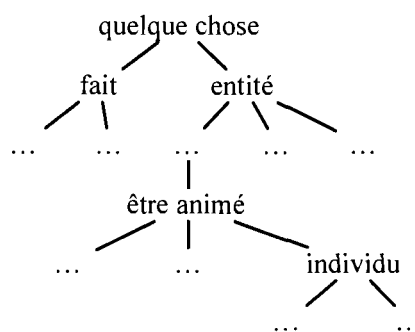


Figure 73 — Un aperçu de la hiérarchie des étiquettes sémantiques du Lexique actif du français

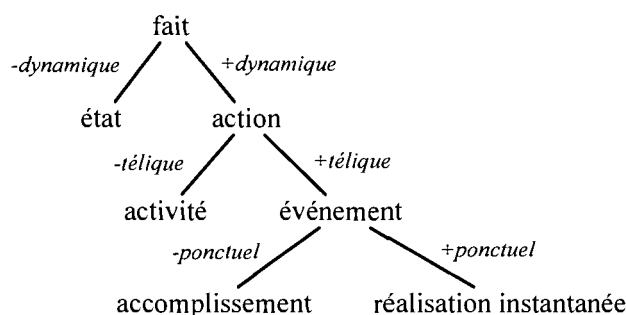


Figure 74 — La classification des faits selon Mourelatos

1.3 Les sens temporels

Nous avons vu au Chapitre VI que les grammèmes de temps en français présentent la particularité de pouvoir situer les faits soit par rapport au moment d'énonciation, soit par rapport à d'autres faits qui peuvent être dans le passé ou dans le futur (*cf.* Ch. VI, § 6, p. 177). Donc, le sens 'maintenant' (qui représente le moment d'énonciation) ne doit pas faire partie de leur définition, mais doit pouvoir en être un actant. En fait, les grammèmes de temps en français n'expriment rien d'autre que l'antériorité, la postériorité ou la simultanéité, ce que nous représentons par les sémantèmes 'avant', 'après' et 'simultané' à la Figure 75, p. 287. Nous considérons ces grammèmes comme sémantiquement indécomposables. Pour les définir, il faut sortir du système de la langue, comme le fait par exemple Gosselin (1996) (*cf.* Ch. V, § 5, p. 106). Dans son modèle, les points *E*, *R* et *S* reichenbachiens⁷ sont

6. Pour une discussion détaillée de la classification de Vendler et de ses critiques, voir Vetter (1996: pp. 87–106).

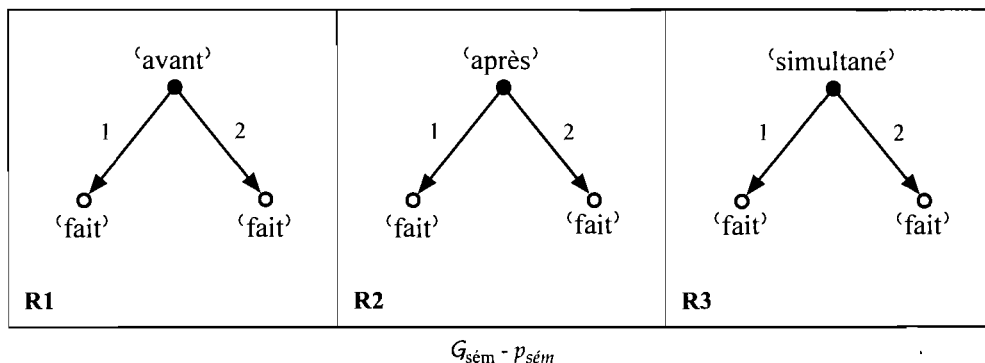


Figure 75 — Les sémantèmes 'avant', 'après' et 'simultané' (R1 à R3)

représentés par des intervalles, respectivement $[B1, B2]$, $[I, II]$ et $[01, 02]$ ⁸. Toutes les relations temporelles entre ces intervalles sont décrites par des combinaisons plus ou moins complexes de trois relations de base entre leurs bornes :

- $i = j$ (i et j coïncident exactement).
- $i \propto j$ (i précède j mais en est infiniment proche).
- $i \not\propto j$ (i précède j et n'est pas dans son voisinage immédiat).

À partir de ces relations de base, il définit quatre relations complexes :

- $i < j \equiv (i \propto j) \vee (i \not\propto j)$
- $i > j \equiv j < i$
- $i \leq j \equiv (i < j) \vee (i = j)$
- $i \geq j \equiv j \leq i$

Ces relations entre bornes servent à définir les relations entre les intervalles :

- $[i, j]$ ANT $[k, l] \equiv j < k$ (tout l'intervalle $[i, j]$ est avant la borne k).
- $[i, j]$ POST $[k, l] \equiv i > l$ (tout l'intervalle $[i, j]$ est après la borne l).
- $[i, j]$ SIMUL $[k, l] \equiv (i \leq l) \& (j \geq k)$ (l'intersection n'est pas vide; les intervalles peuvent coïncider ou se chevaucher).
- $[i, j]$ CO $[k, l] \equiv (i = k) \& (j = l)$ (les intervalles coïncident exactement).
- $[i, j]$ RE $[k, l] \equiv (i < k) \& (j > l)$ ($[k, l]$ est entièrement contenu dans $[i, j]$).
- $[i, j]$ ACCESS $[k, l] \equiv (i \leq k) \& (j \geq l)$ (il y a recouvrement ou coïncidence).

7. Respectivement, le temps du fait dénoté par le verbe, le moment de référence et le moment d'énonciation (cf. Ch. VI, § 6.1.4, p. 187).

8. Pour tout intervalle $[i, j]$, i précède toujours j , même si la distance entre les deux est infiniment petite.

Notre sémantème ‘avant’ correspond exactement à la relation ANT de Gosselin. D’après ses définitions, la relation POST est un simple conversif de ANT, ce qui revient à dire que ‘X est après Y’ \equiv ‘Y est avant X’. D’un point de vue extra-linguistique, c’est tout à fait correct. Cependant, nous croyons qu’il existe dans la langue une différence communicative entre ‘avant’ et ‘après’. Le modèle de Gosselin ne permet pas d’encoder cette différence. Le nôtre ne le permet pas directement non plus, mais cette information se trouve encodée de façon indirecte par la numérotation des actants. En effet, l’ordre des actants d’un prédicat sémantique indique leur saillance communicative relative (*cf.* Mel’čuk 2004).

Notre sémantème ‘simultané’ correspond à la fois aux relations SIMUL, CO, RE et ACCESS de Gosselin. Il ne dénote pas forcément une coïncidence exacte (en fait, il est assez rare que cela se produise). Il faut l’interpréter comme dénotant une «simultanéité subjective».

Nos trois sémantèmes temporels ont la même structure actantielle: leur premier actant est le fait à situer dans le temps; leur second est le point de repère par rapport auquel est situé ce fait. Comme nous l’avons déjà dit à plusieurs reprises, ce point de repère, en français, peut être n’importe quel fait, qu’il s’agisse du moment d’énonciation ou d’un fait dénoté par un autre sémantème. Par exemple, dans la phrase *La branche casse*, le fait dénoté par ‘casser’ est situé directement par rapport au moment d’énonciation, comme le montre la Figure 76a ci-dessous. Par contre, dans *Il m’a dit que la branche casserait*, ‘casser’ est situé par rapport à ‘dire’, comme on peut le voir à la Figure 76b.

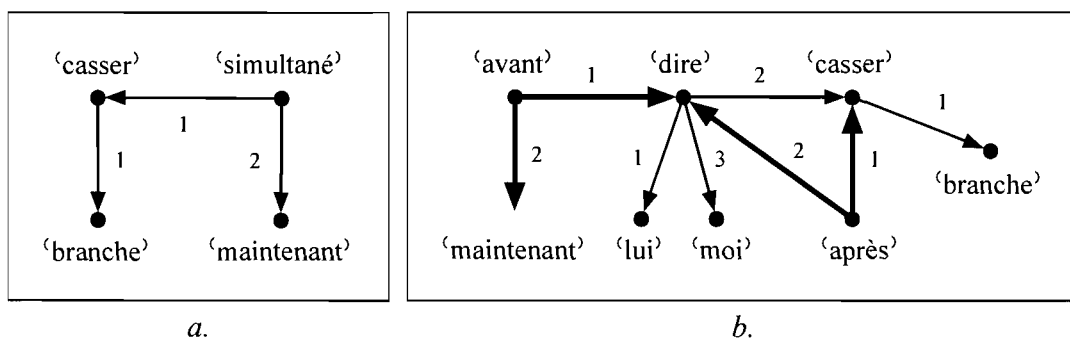


Figure 76 — Les sémantèmes temporels en contexte

Il faut noter que ‘casser’ est quand même lié indirectement à ‘maintenant’ dans la Figure 76b (nous avons mis les arcs en gras pour souligner notre propos), sans qu’il ne soit pour autant possible d’en inférer une relation temporelle entre ‘casser’ et ‘maintenant’. Nous avons vu au Chapitre VI (§ 6.4.3, p. 206) que quand un sémantème est situé temporellement par rapport à un fait dans le passé, comme c’est le cas ici pour ‘casser’, alors le lexème correspondant doit porter le grammème **décalé** (c’est pourquoi on a *casserait*, et non *cassera*). Sinon, c’est le grammème **non-décalé** qui doit être utilisé (c’est pourquoi on a *a dit*, et non *avait dit* pour 76b, ou encore *casse*, et non *cassait* pour 76a). Mais les grammèmes **décalé** et **non-décalé** n’expriment aucun sémantème en soi. Les sémantèmes ‘après’ et ‘avant’ de la Figure 76b et ‘simultané’ de la Figure 76a correspondent respectivement aux grammèmes **postérieur**, **antérieur** et **simultané**; tout ce qu’expriment les grammèmes **décalé** et **non-décalé**, c’est le fait que ‘casser’ (ou ‘dire’) est situé par rapport à un repère passé ou non. Nous verrons plus loin comment cela se modélise dans la grammaire d’interface sémantique-syntaxe (§ 3.4.2, p. 335).

1.4 Les sens phasiques / aspectuels

Nous avons donné au Chapitre VI (§ 5, p. 164) des définitions provisoires pour AVOIR_{acc}, ALLER, ‘EN TRAIN’, VENIR et ‘SUR LE POINT’, que nous répétons ici :

- ‘avoir_{acc} X-é’ = ‘la période qui suit le fait X a cours’
- ‘aller X-er’ = ‘la période qui précède le fait X a cours’
- ‘en train de X-er’ = ‘la période simultanée au fait X a cours’
- ‘venir de X-er’ = ‘la période qui suit immédiatement le fait X a cours’
- ‘sur le point de X-er’ = ‘la période qui précède immédiatement le fait X a cours’

Dans ces définitions, «avoir cours» est vide de sens et ne sert qu’à les rendre plus lisibles; cette expression ne correspond à aucun sémantème dans les définitions formelles.

Nous avons également démontré au Chapitre VI que les sens exprimés par les phases aspectuelles sont mono-actantiels. En effet, on peut observer que les marqueurs de phase sont compatibles avec les verbes à sujet explétif :

(156) *Il a plu : ma selle de vélo est toute mouillée.*

(157) *Il va faire chaud.*

Certaines combinaisons par contre sont agrammaticales :

(158)**Il est sur le point de faire chaud.*

(159)**Il vient de faire un temps humide.*

Ce n'est cependant pas à cause d'un actant manquant, comme en témoignent les phrases suivantes, où l'actant est exprimé :

(160)**La serviette est sur le point d'être jaune.*

(161)**Nous venons d'être croyants.*

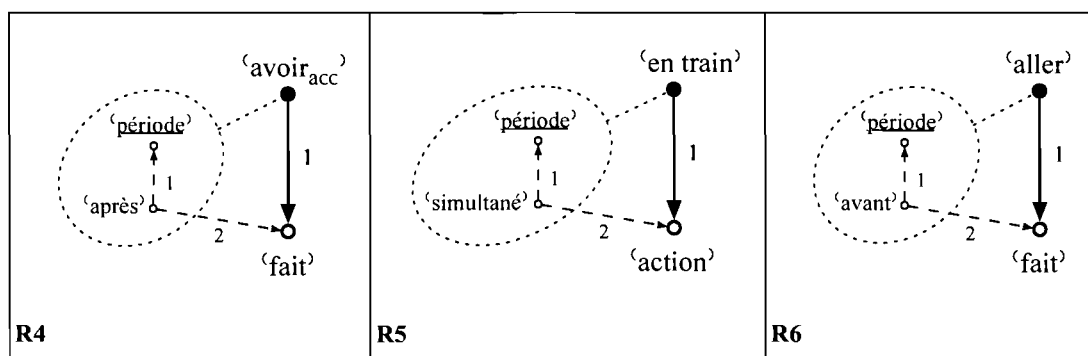
C'est plutôt la nature stative des sens dénotés par les verbes en (158)–(161) qui est incompatible avec ces marqueurs de phase. Seuls AVOIR_{acc} et ALLER sont compatibles avec tous les types de verbes, alors que les autres marqueurs de phase aspectuelle ne peuvent pas se combiner aux états, comme on peut le voir dans le tableau ci-dessous, qui résume la combinatoire sémantique des marqueurs de phase aspectuelle :

	États	Activités	Accomplissements	Réalisations instantanées
AVOIR	<i>Il a déjà eu la foi</i>	<i>Il a mangé</i>	<i>Il a écrit une lettre</i>	<i>Il a trouvé sa clé</i>
ALLER	<i>Il va avoir la foi</i>	<i>Il va manger</i>	<i>Il va écrire une lettre</i>	<i>Il va trouver sa clé</i>
「EN TRAIN」	* <i>Il est en train d'avoir la foi</i>	<i>Il est en train de manger</i>	<i>Il est en train d'écrire une lettre</i>	<i>Il est en train de trouver sa clé</i>
VENIR	* <i>Il vient d'avoir la foi</i>	<i>Il vient de manger</i>	<i>Il vient d'écrire une lettre</i>	<i>Il vient de trouver sa clé</i>
「SUR LE POINT」	* <i>Il est sur le point d'avoir la foi</i>	<i>Il est sur le point de manger</i>	<i>Il est sur le point d'écrire une lettre</i>	<i>Il est sur le point de trouver sa clé</i>

Tableau XXVII — La combinatoire sémantique des marqueurs de phase aspectuelle

On peut rendre compte du phénomène par un typage approprié de l'actant des sémantèmes phasiques: 'avoir_{acc}' et 'aller' prennent comme argument un fait, alors que les autres prennent une action (nous avons vu plus haut, Figure 74, p. 286, que la classification de Mourelatos regroupe tous les faits non statifs sous l'étiquette «action»).

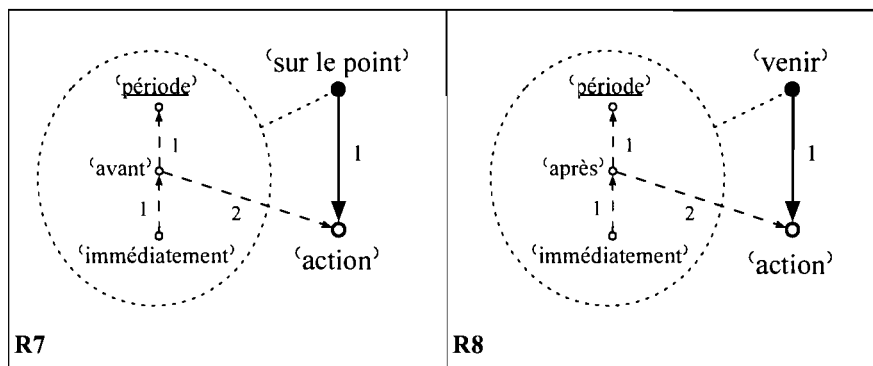
Nous représentons le sens des signes de phase aspectuelle par un sémantème étiqueté du marqueur qui l'exprime. Trois des signes de phase aspectuelles dénotent chacune des trois phases naturelles délimitées par les bornes d'un fait (*cf.* Ch. VI, § 5.1, p. 165):



$G_{\text{sém}} - P_{\text{sém}}$

Figure 77 — Les sémantèmes 'avoir_{acc}', 'en train' et 'aller' (R4 à R6)

Les deux autres signes dénotent la période immédiatement avant le début ou immédiatement après la fin d'une action:



$G_{\text{sém}} - P_{\text{sém}}$

Figure 78 — Les sémantèmes 'sur le point' et 'venir' (R7 et R8)

Les sémantèmes de phase se comportent comme n'importe quel sémantème dénotant un fait. Ils peuvent prendre comme actant n'importe quel fait (ou action, selon leur combinatoire), y compris un autre sémantème de phase aspectuelle. La Figure 79a montre la représentation sémantique de *Ils ont bien dormi*, alors que la Figure 79b donne celle de *Ils vont avoir bien dormi*.

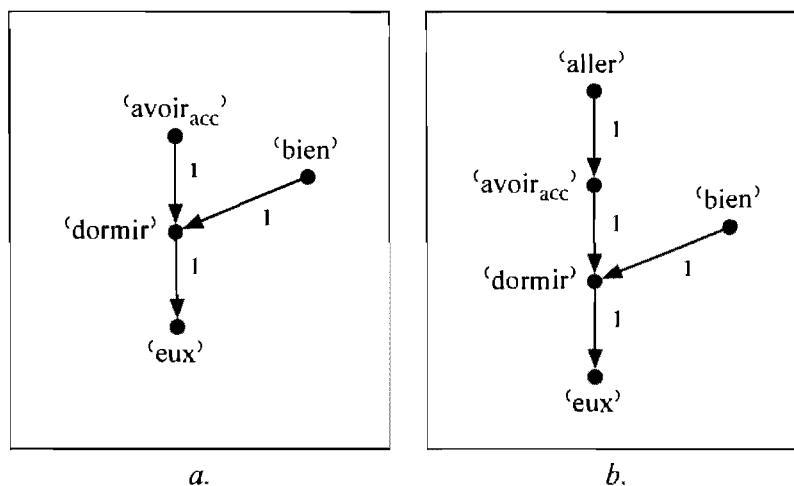


Figure 79 — Les sémantèmes de phase aspectuelle en contexte

On peut représenter schématiquement la situation dénotée par *Ils vont avoir bien dormi* comme suit (la position de T_0 — le moment d'énonciation — par rapport au moment du sommeil n'est pas pertinente, ce que nous représentons ici par deux points T_0 concurrents):

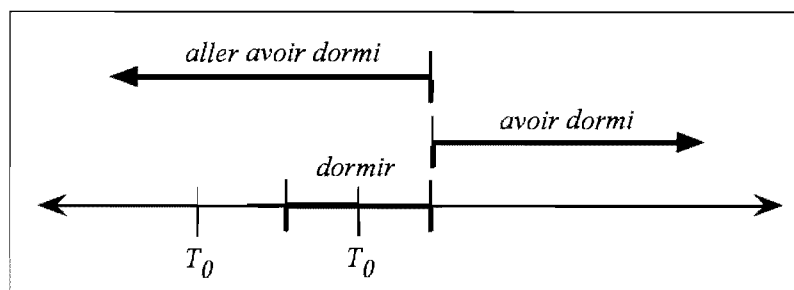


Figure 80 — La situation dénotée par *Ils vont avoir bien dormi*

Toutes les combinaisons ne sont pas possibles, par contre. Par exemple, puisque 'en train' demande un actant de type 'action', aucun des autres sémantèmes phasiques ne pourront être actant de 'en train', leur genre prochain ('période') étant

incompatible. Des contraintes syntaxiques sont également à l'œuvre (cf. § 2.3.1, p. 303).

Nous ne voyons pas de raison pour que les sémantèmes de phase aspectuelle ne puissent aussi être des actants d'autres types de prédicats. Par exemple, nous représentons le sens de la phrase *Je pense qu'ils sont en train de dormir* par la structure de la Figure 81, où l'actant de 'penser' n'est pas 'dormir', mais 'en train'.

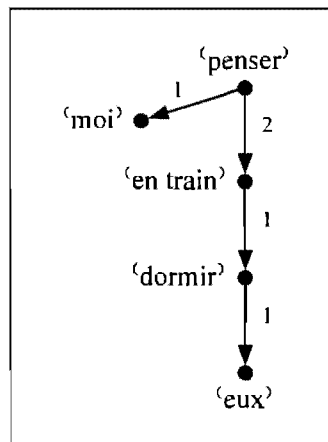


Figure 81 — Un sémantème de phase aspectuelle actant de 'penser'

Il découle de ce que nous venons de dire que les sémantèmes de phase aspectuelle peuvent être situés dans le temps comme n'importe quel autre fait. La Figure 82 ci-dessous (page suivante) reprend la représentation sémantique de la Figure 79b en y ajoutant la situation temporelle ainsi que les décompositions sémantiques de 'aller' et 'avoir_{acc}'. On voit que la phrase *Ils vont avoir bien dormi* signifie 'c'est maintenant la période avant la période qui suit eux-bien-dormir' (cf. Figure 80, p. 292).

Ici, nous avons appliqué la règle de réancrage des gouverneurs que nous avons présentée au Chapitre VII (§ 6, p. 267), et que nous répétons en tant que R9 ci-dessous (page suivante) puisqu'elle fait partie de notre modèle.

Enfin, nous avons vu au Chapitre VI (§ 1.2.2, p. 124) que le participe passé peut avoir une valeur résultative [*Nous sommes arrivés*]. Il dénote alors l'état résultant d'un fait. Contrairement aux signes de phase aspectuelle, le sémantème en jeu ici

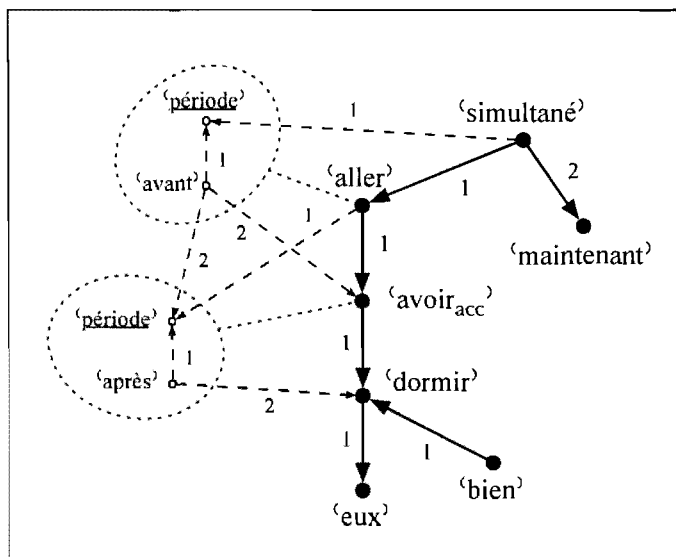
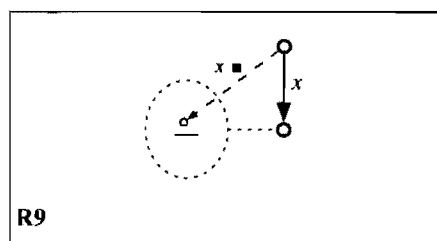


Figure 82 — L'interaction sémantique des temps et des phases aspectuelles



$G_{sém} - P_{sém} (P_{sém-syn})$

Figure 83 — Le réancrage des gouverneurs sémantiques (R9)

(que nous appellerons *résultatif*⁹) possède deux actants. On peut le vérifier en observant l'agrammaticalité de la construction résultative d'un verbe impersonnel, comme en (162) ci-dessous.

(162)**Il est plu.*

Cela est dû au fait que le résultatif dénote l'état de quelque chose; ce quelque chose doit être un actant de *résultatif*. Par ailleurs, son second actant ne peut pas être n'importe quel type de fait:

(163)**Il est lévité. (état)*

(164)[#]*Du pain est mangé. (activité)*¹⁰

9. Nous le mettons en italiques pour indiquer qu'il s'agit d'une étiquette métalinguistique. En effet, ce sens n'est pas celui de la *lexie* RÉSULTATIF.

10. Il ne faut pas confondre avec l'interprétation passive de cet énoncé.

(165) *La lettre est écrite.* (accomplissement)

(166) *Tous nos objectifs sont atteints.* (réalisation instantanée)

Ce ne sont que les événements (dans la classification de Mourelatos 1978) qui peuvent être le deuxième actant de *'résultatif'*. La propriété distinctive des événements étant leur télélicité (*cf.* Ch. VI, § 6.1.2, p. 181), cela nous indique que *'résultatif'* inclut dans son sens l'idée d'achèvement d'un fait télélique.

Donc, nous pouvons définir le sens du participe passé résultatif ainsi : 'X résultatif de l'événement Y' \equiv 'état de X causé par l'achèvement de l'événement Y'. Formellement, cela se représente comme suit :

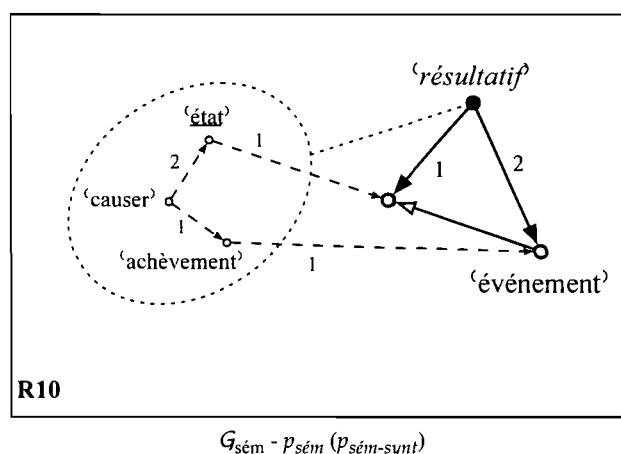


Figure 84 — Le sémantème *'résultatif'* (R10)

1.5 Les sens modaux

Les sens «modaux» (dans une acception large du terme) sont d'un intérêt limité pour notre thèse. Nous nous contenterons d'en donner une description sémantique sommaire, sans en fournir la décomposition.

La finitude est essentiellement une catégorie flexionnelle syntaxique, c'est-à-dire que ses grammèmes n'expriment pas un sens, mais signalent une configuration syntaxique (*cf.* Ch. VI, § 1, p. 116). Nous avons vu néanmoins que deux de ces grammèmes peuvent être porteurs d'un sens modal.

Le grammème profond **gérondif** exprime la «manière» au sens large. Ce sens englobe le moyen de réalisation [*Il ouvre la bouteille en tirant le bouchon*], la caractérisation [*Il marche en boitant*] et la simultanéité [*Il danse en chantant*]. Nous le représentons par le sémantème 'manière', un prédicat binaire dont le premier argument est le fait à caractériser et dont le second est la caractérisation de ce fait :

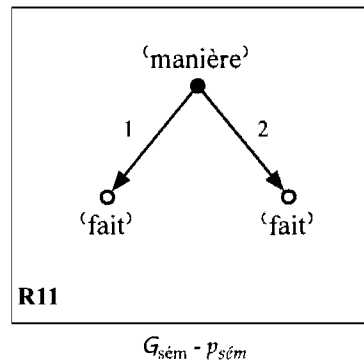


Figure 85 — Le sémantème 'manière' (R11)

Le grammème **infinitif**, quand il est le sommet syntaxique de la phrase, exprime une consigne. Le sens qui lui correspond est le prédicat 'devoir', que nous représentons par la règle R12 ci-dessous.

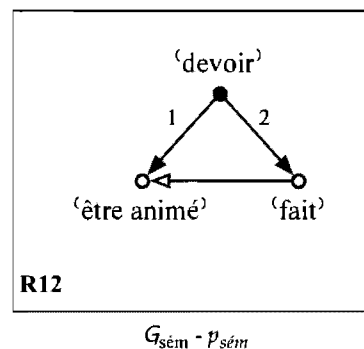


Figure 86 — Le sémantème 'devoir' (R12)

Cette règle décrit le prédicat 'devoir' non pas en tant que signifié de l'infinitif de consigne, mais en tant que tel. Nous croyons que le sens de l'infinitif de consigne se distingue légèrement de celui de l'impératif, pourtant très proche. Ce dernier implique directement le locuteur (ce qui expliquerait qu'on lui préfère l'infinitif pour les consignes écrites). Il exprime le sémantème 'ordonner', que décrit la règle

R13 ci-dessous. Encore une fois, cette règle ne décrit pas ce prédicat en tant que signifié de l'impératif, mais en lui-même.

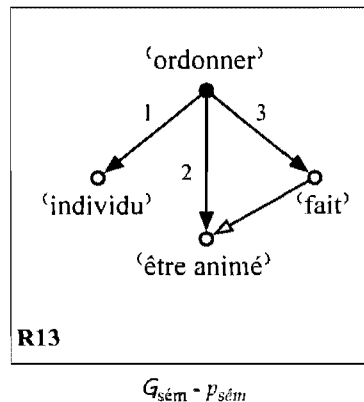


Figure 87 — Le sémantème 'ordonner' (R13)

Nous avons également vu au Chapitre VI (§ 6.4.4, p. 213) trois phrasèmes grammémiques qui ont un sens modal. Le premier est le futur de supposition [*Mais où sont-ils? Ils se seront sans doute perdus en chemin*], dont nous représentons le sens par le sémantème 'supposer' dans la règle R14 ci-dessous.

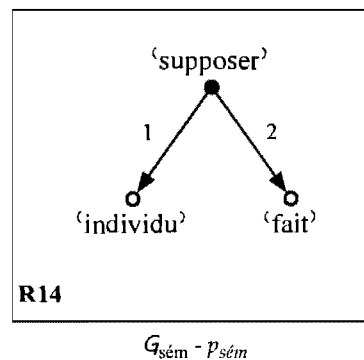


Figure 88 — Le sémantème 'supposer' (R14)

Le second est le conditionnel de réserve [*Le président de syndicat aurait été arrêté hier soir*]. Nous en représentons le sens par 'il paraît' à la règle R15 ci-dessous.

Enfin, le dernier est le futur antérieur de rétrospective [*L'hiver aura été particulièrement long cette année*], dont nous représentons le sens par le sémantème 'en rétrospective' à la règle R16.

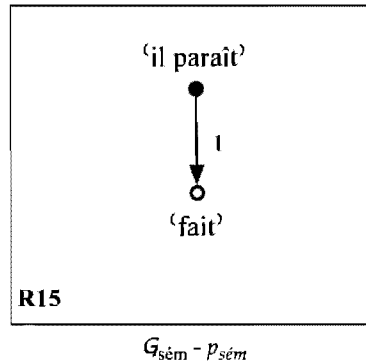


Figure 89 — Le sémantème 'il paraît' (R15)

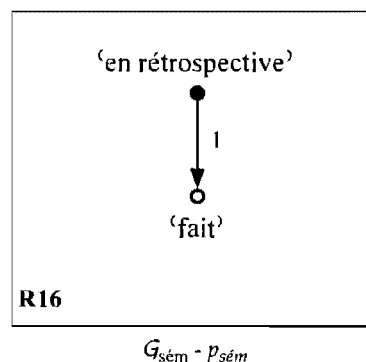


Figure 90 — Le sémantème 'en rétrospective' (R16)

2 La grammaire de bonne formation syntaxique

La grammaire de bonne formation syntaxique, G_{synt} , sert à assurer la construction d'arbres syntaxiques valides. Deux règles, regroupées dans le sous-module G_{arbre} , suffisent à vérifier la bonne formation structurelle des arbres syntaxiques. D'autres règles doivent décrire la combinatoire des éléments des représentations syntaxiques, par exemple les relations syntaxiques valides dans la langue décrite. En principe, les règles du module d'interface syntaxe-morphotopologie doivent décrire la combinatoire morphologique des lexèmes et des grammèmes. Comme nous n'avons pas ce module dans notre modèle, nous projetons cette information dans G_{synt} .

Ainsi, il existe quatre types de règles en G_{synt} : les règles de bonne formation structurelle (G_{arbre}) ainsi que les règles lexicales, sagittales et grammémiques.

Nous les présenterons dans cet ordre; mais d'abord, rappelons les caractéristiques formelles de $\mathcal{G}_{\text{synt}}$.

2.1 Les caractéristiques formelles de $\mathcal{G}_{\text{synt}}$

Nous rappelons ici les principales caractéristiques formelles de la grammaire de bonne formation syntaxique. Pour une présentation plus complète, le lecteur pourra relire la section correspondante du chapitre précédent (Ch. VII, § 5.1.2, p. 242).

Le rôle de $\mathcal{G}_{\text{synt}}$ est de vérifier la bonne formation des représentations syntaxiques construites par les grammaires d'interface $\mathcal{I}_{\text{sém-synt}}$ (en synthèse) ou $\mathcal{I}_{\text{synt-topo}}$ (en analyse). Cela se fait par la saturation de la polarité p_{synt} laissée en blanc par ces modules. $\mathcal{G}_{\text{synt}}$ utilise également une polarité p_{arbre} qui sert à vérifier la bonne formation structurelle de l'arbre syntaxique. Cette dernière est saturée par des règles que nous regroupons dans un sous-module de $\mathcal{G}_{\text{synt}}$, que nous appelons $\mathcal{G}_{\text{arbre}}$ (cf. § 2.2, p. 300).

Les objets de $\mathcal{G}_{\text{synt}}$ sont de trois types: nœuds syntaxiques (qui représentent les lexèmes), arcs syntaxiques (qui représentent les relations syntaxiques) et objets grammémiques (qui représentent les grammèmes superficiels). Tous doivent porter une polarité p_{synt} . Les nœuds et les arcs portent également une polarité p_{arbre} , qui est toujours blanche sauf dans les règles de $\mathcal{G}_{\text{arbre}}$, où c'est plutôt p_{synt} qui est blanche. Par souci de lisibilité, la polarité p_{arbre} sera donc systématiquement omise dans les figures, sauf bien entendu dans les règles de $\mathcal{G}_{\text{arbre}}$. Par ailleurs, tous les objets, y compris les objets grammémiques, portent deux polarités d'articulation: $p_{\text{sém-synt}}$ et $p_{\text{synt-topo}}$. Ces dernières sont la plupart du temps blanches, aussi les omettons-nous dans presque toutes les figures. Nous les représentons seulement lorsqu'elles sont saturées.

Les objets grammémiques apparaissent comme des losanges dans les représentations graphiques. Ils ne doivent pas être confondus avec les traits de syntactique. Les objets grammémiques sont des objets des structures GUST au même titre que

les nœuds et les arcs syntaxiques. Ils doivent donc être polarisés. Ils représentent les grammèmes superficiels qui sont associés aux lexèmes. Les traits de syntactique¹¹, quant à eux, sont des propriétés intrinsèques des lexèmes. Ils ne sont pas représentés par des objets dans les structures GUST, mais plutôt par des fonctions d'étiquetage (*cf.* Ch. VII, § 2, p. 226).

Par exemple, dans la règle R19 de la Figure 92 (p. 302), on trouve trois traits de syntactique pour le lexème RÊVER: sa partie du discours, l'auxiliaire auquel il s'associe et son groupe de conjugaison. Ces propriétés de RÊVER ne changent jamais¹², peu importe le contexte. Elles sont représentées par des fonctions dont la valeur n'est qu'une étiquette, et non un objet de la structure (d'où l'absence de polarité). Par contraste, la fonction *fin(itude)*¹³, dans la même règle, retourne comme valeur un objet grammémique polarisé en blanc. Cet objet n'est pas étiqueté, mais il ne faut pas croire pour autant que la fonction *fin* n'a pas de valeur dans cette structure. Sa valeur, c'est l'objet grammémique lui-même, et non l'étiquette de cet objet. Que cet objet soit étiqueté ou non n'a aucune importance du point de vue de *fin*: cette fonction ne fait qu'associer un nœud syntaxique à un objet grammémique.

2.2 La grammaire d'arbres

Nous avons vu au Chapitre VII (§ 5.1.2, p. 242) que G_{synt} devait être articulé avec une grammaire d'arbres, G_{arbre} . Il s'agit en fait d'un sous-module de G_{synt} dont les règles se distinguent par la saturation des objets en p_{arbre} . Son rôle est de vérifier que les structures syntaxiques construites par G_{synt} sont bien des arbres. Pour ce faire, deux règles suffisent. La règle R17 ci-dessous constitue la structure initiale de cette grammaire (*cf.* Ch. VII, § 2, p. 226), ce qu'indique le double enca-

11. On les appelle également «caractéristiques grammaticales» ou «propriétés de combinatoire».

12. Même si RÊVER se trouve avec l'auxiliaire ÊTRE dans *Ils s'étaient révé une vie pas possible*, il reste que sa propriété intrinsèque est de sélectionner AVOIR. C'est la construction réfléchie qui impose ici l'auxiliaire ÊTRE, peu importe la combinatoire du verbe auxilié.

13. Rappelons qu'il s'agit d'une fonction structurante (*cf.* Ch. VII, § 2, p. 226).

dré. Rappelons qu'une structure initiale doit être utilisée une et une seule fois. Cette règle construit le sommet de l'arbre. La règle R18 en construit tous les dépendants¹⁴.

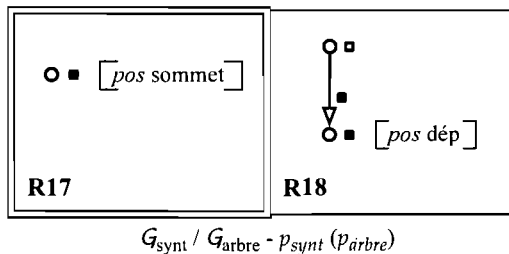


Figure 91 — La grammaire d'arbres (R17 et R18)

Nous marquons tous les nœuds d'un trait *pos(ition)*, dont la valeur peut être «sommet» ou «dép(endant)». Ce trait pourra être utilisé dans les règles pour vérifier qu'un nœud donné est ou non le sommet de l'arbre¹⁵.

2.3 Les règles lexicales

Les règles lexicales de G_{synt} donnent l'inventaire des lexèmes de la langue et décrivent leur combinatoire morphologique (cette information est en fait une projection de la grammaire d'interface syntaxe-morphotopologie). Ainsi, toute règle introduisant un verbe doit en fournir les traits de syntactique, tels son groupe de conjugaison et l'auxiliaire auquel il peut se combiner (*AVOIR* ou *ÊTRE*), comme la règle R19 ci-dessous¹⁶.

Les règles lexicales verbales doivent également introduire un objet grammématique polarisé en blanc pour forcer l'attribution d'une finitude au verbe. En un sens, cela est dangereux puisque le formalisme ne contraint pas le linguiste à toujours mettre cet objet dans les règles qui construisent les nœuds verbaux. L'oublier aurait pour conséquence d'autoriser un verbe à apparaître sans aucune flexion (les autres

14. Afin de garder une certaine homogénéité par rapport aux autres figures, nous avons choisi de montrer les règles de G_{arbre} avec p_{synt} comme polarité principale. La polarité p_{arbre} apparaît comme polarité secondaire.

15. Voir les règles R66 (p. 322) et R89 (p. 343).

16. Par souci de lisibilité, nous omettons ces traits dans les autres figures, sauf là où ils sont pertinents pour notre propos.

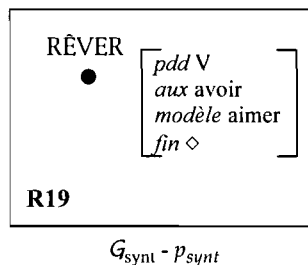


Figure 92 — Un règle demandant une finitude pour un verbe (R19)

catégories flexionnelles étant dépendantes de la finitude). Dans le cadre d'une grammaire concrète à large couverture, l'erreur semble inévitable. Il y a toutefois au moins deux solutions envisageables à ce problème.

La première solution consiste à représenter la partie du discours par un objet. En polarisant cet objet en blanc pour tous les verbes, on forcerait l'application d'une règle générale qui ne ferait que saturer cet objet tout en introduisant des objets grammémiques en blanc pour représenter les grammèmes superficiels nécessaires à tous les éléments de la partie du discours en question.

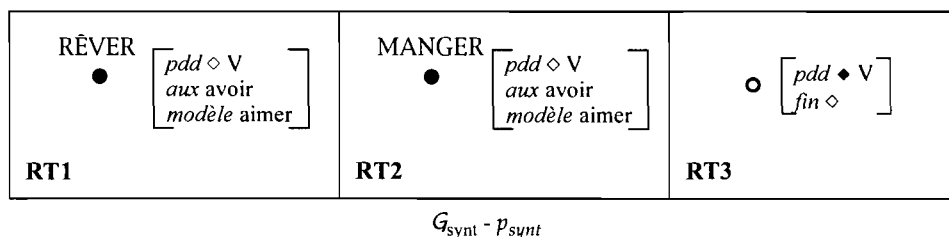


Figure 93 — Une représentation alternative des parties du discours (RT1 à RT3)

Cette méthode présente l'avantage de faciliter énormément l'entretien de la grammaire. Si, après avoir entré les règles pour RÉVER, MANGER et quelques centaines d'autres verbes, on réalise qu'il ne faut finalement non seulement associer à chaque verbe une finitude, mais également une autre catégorie flexionnelle, on n'a alors qu'une seule règle à modifier, RT3.

Cependant, nous préférons ne pas représenter les traits de syntactique par des objets polarisés puisqu'il s'agit de propriétés «inertes». La polarisation des objets sert à contrôler la composition des signes, un processus en un sens plus «dynamique» que les propriétés de combinatoire des signes linguistiques. Nous

préférons que notre formalisme sépare bien les deux phénomènes en les représentant de façon distincte.

Une deuxième solution consiste à utiliser des patrons de règles. Il suffit alors d'avoir un patron pour les verbes dans lequel apparaît l'objet grammémique blanc pour la finitude, et cette partie de la structure sera héritée par toutes les règles marquées comme héritant du patron «verbe». En d'autres termes, il faut avoir pour chaque langue des règles générales qui stipulent quelles sont les catégories flexionnelles pour chaque partie du discours. Ce type de mécanisme est couramment utilisé dans d'autres cadres théoriques.

On peut en outre faire appel à des outils d'édition appropriés (qui, de toute façon, s'avèrent nécessaires dès qu'on envisage un modèle à couverture assez large). En fait, le linguiste ne devrait jamais avoir à introduire toutes les règles lexicales manuellement; il devrait suffire de mettre dans un dictionnaire que RÊVER est un verbe pour en dériver automatiquement la règle R19 (*cf.* Lareau 2002).

Les règles lexicales qui nous intéressent particulièrement sont celles qui décrivent la combinatoire des auxiliaires et semi-auxiliaires du français.

2.3.1 *La combinatoire des marqueurs de phase aspectuelle*

Les seuls véritables auxiliaires de phase aspectuelle sont ALLER, VENIR, AVOIR_{acc} et ÊTRE_{acc}. Nous avons vu au tout début de cette thèse que les autres constructions, notamment 「EN TRAIN」 et 「SUR LE POINT」, n'étaient pas des auxiliaires (Ch. II, § 3, p. 28). Il s'agit en effet de locutions prépositionnelles et leur traitement ne diffère pas de celui des autres locutions prépositionnelles. Au niveau syntaxique, la représentation de 「EN TRAIN」 et 「SUR LE POINT」 n'est pas structurellement différente de celle, par exemple, de 「À CÔTÉ」, comme on peut le voir à la Figure 94, p. 304. Il n'est pas nécessaire d'avoir des règles spécifiques pour ces locutions, puisque leur représentation syntaxique peut être entièrement construite à partir des règles de $\mathcal{G}_{\text{synt}}$ qui décrivent les lexèmes et les arcs qui les constituent.

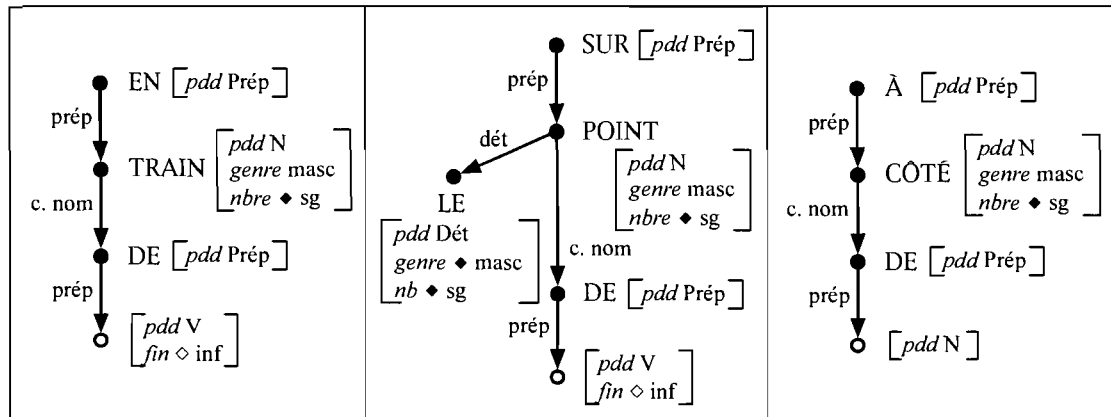


Figure 94 — La structure syntaxique de trois locutions prépositionnelles

Ces règles sont triviales et il y a peu d'intérêt à toutes les énumérer ici. Nous en donnons quelques-unes à titre illustratif ci-dessous.

● À [pdd Prép]	● EN [pdd Prép]	● CÔTÉ [pdd N genre masc nbre ♦]
R20	R21	R22

$G_{\text{synt}} - p_{\text{synt}}$

Figure 95 — Quelques règles lexicales triviales (R20 à R22)

Revenons maintenant aux vrais auxiliaires. Les marqueurs de l'accompli peuvent accepter n'importe quelle finitude, mode, etc. Par contre, ALLER et VENIR ont une combinatoire plus capricieuse (le Tableau I, p. 30, résume la combinatoire morphologique des auxiliaires). Les règles R23 à R28 ci-dessous modélisent la combinatoire de ces lexèmes.

La combinatoire de VENIR étant particulièrement capricieuse, elle ne se laisse pas bien décrire par des règles assez générales. C'est pourquoi nous avons trois règles pour cet auxiliaire. Remarquons qu'au contraire, bien que ALLER ne se combine pas librement à n'importe quel grammème non plus (il ne peut apparaître qu'au présent ou à l'imparfait), une seule règle suffit à en décrire la combinatoire. Cela est dû au fait que le présent et l'imparfait, selon notre description, ne se distinguent que par le grammème de décalage (**non-décalé** pour le présent et **décalé**

AVOIR _{acc} ● $\left[\begin{array}{l} pdd \vee \\ fin \diamond \end{array} \right]$ R23	ÊTRE _{acc} ● $\left[\begin{array}{l} pdd \vee \\ fin \diamond \end{array} \right]$ R24	ALLER ● $\left[\begin{array}{l} pdd \vee \\ fin \diamond fini \\ mode \diamond ind \\ temps \diamond sim \end{array} \right]$ R25
VENIR ● $\left[\begin{array}{l} pdd \vee \\ mode \diamond ind \end{array} \right]$ R26	VENIR ● $\left[\begin{array}{l} pdd \vee \\ mode \diamond subj \end{array} \right]$ R27	VENIR ● $\left[\begin{array}{l} pdd \vee \\ fin \diamond part-ant \end{array} \right]$ R28

 $G_{synt} - p_{synt}$

Figure 96 — La combinatoire des auxiliaires de phase aspectuelle (R23 à R28) pour l'imparfait). Dans un modèle où temps et décalage ne seraient pas distingués, il faudrait deux règles pour ALLER.

Les signes de phase aspectuelle peuvent se combiner entre eux. Toutefois, l'ordre des auxiliaires dans la structure syntaxique n'est pas libre. Par exemple, ALLER peut gouverner VENIR ou AVOIR_{acc} [*Il va bientôt avoir fini*], mais pas l'inverse [*#Il a été finir*]. Le bon ordre structurel des auxiliaires est assuré par leur combinatoire. En effet, ALLER ne peut pas dépendre de VENIR puisque VENIR impose l'infinitif à son dépendant (cf. R79, p. 332), alors que ALLER n'est pas compatible avec cette finitude. Des contraintes sémantiques contribuent aussi à gérer la combinaison des auxiliaires (cf. § 1.4, p. 289).

2.3.2 La combinatoire des auxiliaires de temps

Le seul grammème profond de temps à s'exprimer par un auxiliaire est **antérieur**. Ce dernier varie en fonction du verbe qui en dépend. Contrairement aux marqueurs de l'accompli, les auxiliaires temporels ont une combinatoire relativement restreinte: ils ne peuvent apparaître qu'au présent ou à l'imparfait de l'indicatif, c'est-à-dire qu'ils doivent porter le grammème **simultané**. En ce sens, leur combinatoire est semblable à celle de l'auxiliaire ALLER (cf. R25). Par contre, le grammème **simultané** dans les règles R29 et R30 n'exprime aucun sens en soi, alors que celui de la règle R25 exprime son sens habituel. En effet, dans une phrase

au passé, l'auxiliaire dénote précisément l'antériorité, et non une situation qui peut être présentée comme simultanée à quelque point de référence, alors que ALLER dénote toujours un état simultané à un point de référence (qui peut être aussi bien dans le présent que dans le passé). Donc, la combinaison ALLER \oplus **simultané** est sémantiquement compositionnelle, alors que ce n'est pas le cas pour AVOIR_{ant} et ÊTRE_{ant}, où **simultané** fait partie du signifiant du passé composé à valeur d'antériorité. Il ne doit pas apparaître dans les règles R29 et R30, mais plutôt dans les règles correspondantes de l'interface sémantique-syntaxe, à savoir R82 et R83 (p. 336).

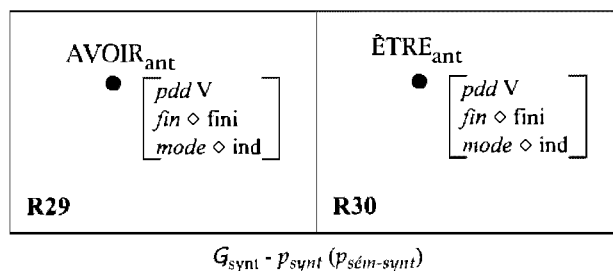


Figure 97 — La combinatoire des auxiliaires de temps (R29 et R30)

2.3.3 Un auxiliaire du passif?

Nous avons vu plus haut (Ch. VI, § 2, p. 132) que le verbe ÊTRE des constructions passives ne présente aucune différence avec la copule. Il n'est donc pas nécessaire de le traiter séparément dans G_{synt} . Il suffit d'avoir une règle qui construit ÊTRE_{cop}, et de s'assurer qu'il peut prendre un actant qui soit un verbe au participe passé (c'est-à-dire un verbe au passif), grâce à une règle sagittale (la règle R53 que nous verrons plus loin, p. 315).

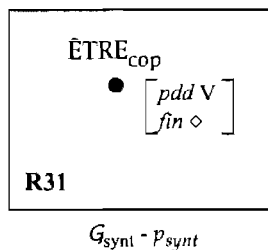


Figure 98 — La règle lexicale pour ÊTRE_{cop} (R31)

2.4 Les règles grammémiques

Les règles grammémiques de G_{synt} donnent la liste des grammèmes superficiels de chaque catégorie flexionnelle, ainsi que leur combinatoire¹⁷.

2.4.1 La combinatoire des grammèmes superficiels de finitude

Nous avons vu que la catégorie flexionnelle de finitude contient cinq grammèmes profonds: **fini**, **infinitif**, **participe-ant**, **particip-é** et **gérondif** (cf. Ch. VI, § 1, p. 116). Seuls les quatre premiers s'expriment de façon synthétique. Ils sont donc aussi des grammèmes superficiels et doivent être décrits dans G_{synt} . Les règles R32 à R34 ci-dessous en décrivent la combinatoire. Ces grammèmes superficiels peuvent saturer les règles lexicales verbales comme R19 (p. 302) ou R23–R31 (pp. 305–306).

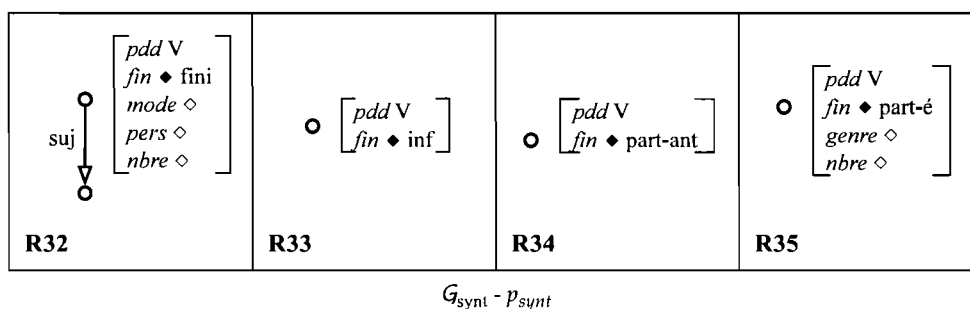


Figure 99 — La combinatoire des grammèmes de finitude (R32 à R35)

Le grammème profond **gérondif** n'a pas besoin d'être décrit en G_{synt} . Il s'exprime par la combinaison de la préposition EN et du grammème superficiel **participe-ant**. Ces éléments sont déjà décrits par d'autres règles¹⁸, dont la combinaison suffit à construire la représentation syntaxique de «*en V-ant*».

Il n'est pas nécessaire d'indiquer ici la diathèse associée à **participe-ant** et **participe-é**. Il ne faut pas perdre de vue deux choses: **1)** ce que nous décrivons ici ne sont que les grammèmes *superficiels*, et **2)** G_{synt} ne sert qu'à modéliser la com-

17. Lorsque nous parlons de la combinatoire des grammèmes d'une catégorie flexionnelle, nous entendons par là la capacité de chacun de ces grammèmes à se combiner avec des grammèmes d'*autres* catégories flexionnelles, et non entre eux, ce qui serait contraire à notre définition du grammème (cf. Ch. IV, § 6, p. 62).

18. À savoir, les règles R21 (p. 304), R34 (p. 307) et R55 (p. 316).

binatoire syntaxique et morphologique des signes; la diathèse et les opérations sur celle-ci relèvent de $I_{\text{sém-synt}}$ puisqu'elles font référence à la structure sémantique (et, dans ce cas-ci, à la structure communicative). Nous y reviendrons plus loin (§ 3.2.3, p. 322).

Le grammème **participe-é** appelle l'accord en genre et en nombre. Nous ne traitons pas ce phénomène dans cette thèse (voir la remarque à ce sujet, § 2.4.5, p. 310).

2.4.2 La combinatoire des grammèmes superficiels de mode

Les verbes finis, comme l'indique la règle R32 (p. 307), demandent un grammème de mode. Les grammèmes superficiels de mode sont décrits par les règles R36 à R39 ci-dessous. On remarquera que l'impératif bloque la réalisation du sujet en neutralisant la polarité $p_{\text{synt-topo}}$ de l'arc et du nœud¹⁹, ce qui est représenté par les pointillés (cf. § 3.3.3, p. 331). C'est cette polarité qui force la mise en correspondance des objets syntaxiques avec des objets du niveau morpho-topologique. En la saturant, on rend donc ces objets inertes pour l'interface. Les règles de l'impératif (R38 et R39) restreignent également les valeurs possibles pour le nombre et la personne (première personne du pluriel, deuxième personne du singulier ou du pluriel).

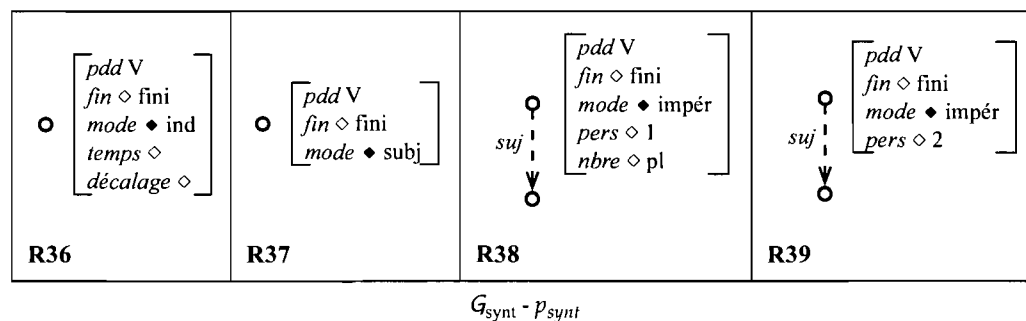


Figure 100 — La combinatoire des grammèmes de mode (R36 à R39)

19. Rappelons que tous les objets de G_{synt} doivent porter, en plus de la polarité p_{synt} , les polarités $p_{\text{sém-synt}}$, $p_{\text{synt-topo}}$ et p_{arbre} . (cf. Ch. VII, § 5, p. 236). Ces polarités sont occultées dans les représentations graphiques lorsque leur valeur est blanche.

C'est d'abord et avant tout l'indicatif qui nous intéresse, puisque ce n'est qu'à ce mode que sont exprimés les grammèmes des catégories flexionnelles de temps et de décalage (*cf.* Ch. VI, § 6, p. 177).

Les règles R36 à R39 n'introduisent pas de grammèmes de personne et de nombre en blanc, sauf pour en contraindre la valeur. Pourtant, les verbes à l'indicatif, au subjonctif ou à l'impératif doivent avoir un nombre et une personne, alors pourquoi ne pas l'indiquer ici? Nous croyons qu'il s'agit en fait d'une propriété de combinatoire du grammème **fini**, et non des grammèmes de mode (*cf.* Figure 29, p. 221). Si les modes apparaissent toujours avec nombre et personne, c'est simplement parce qu'ils ne peuvent se combiner qu'au grammème de finitude **fini**, et que ce dernier exige aussi la présence de grammèmes de personne et de nombre (*cf.* R32). Ces deux dernières catégories sont des catégories d'accord et sont traitées plus loin (§ 2.4.5, p. 310).

Enfin, remarquons qu'il n'y a pas de grammème ***conditionnel**. Il existe indéniablement des signes grammaticaux profonds correspondant aux différentes acceptions du conditionnel de la grammaire traditionnelle, mais il n'existe pas de signifiant qui leur soit propre. Tous s'expriment par la combinaison des grammèmes **postérieur** et **décalé**.

2.4.3 *La combinatoire des grammèmes superficiels de temps*

Dans le registre neutre, le grammème profond **antérieur** s'exprime toujours par un auxiliaire (*cf.* R29 et R30, p. 306) et il n'y a donc que les grammèmes superficiels de temps **simultané** et **postérieur**, qui ne se combinent tous les deux qu'au mode indicatif. Cependant, dans un registre littéraire, le grammème **antérieur** peut également s'exprimer conjointement avec le grammème **non-décalé** dans le morphème du passé simple. Il est donc nécessaire d'avoir au niveau syntaxique un troisième grammème superficiel de temps pour ces formes. La formalisation de la structure rhétorique, où devraient apparaître les contraintes sur le registre, n'ayant pas été étudiée, nous indiquons simplement le registre en langue naturelle dans la règle correspondante à la Figure 101 ci-dessous.

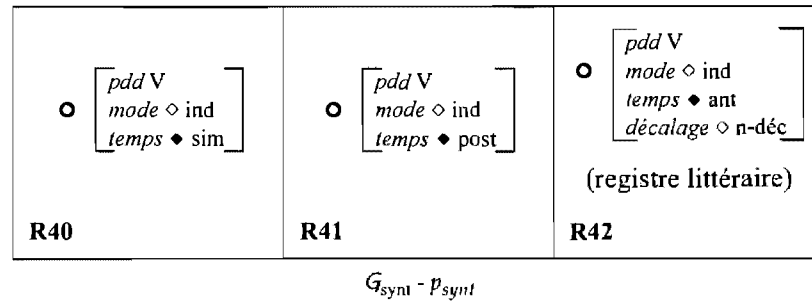


Figure 101 — La combinatoire des grammèmes de temps (R40 et R42)

2.4.4 La combinatoire des grammèmes superficiels de décalage

Comme les grammèmes superficiels de temps, ceux de décalage ne se combinent qu'au mode indicatif.

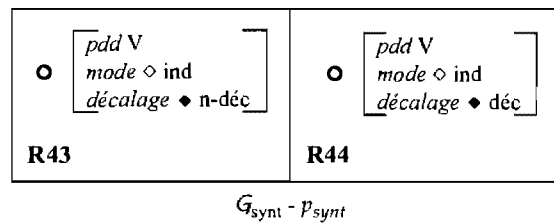


Figure 102 — La combinatoire des grammèmes de décalage (R43 et R44)

Il n'est pas nécessaire (mais il est tout de même possible) d'indiquer dans les règles R40 à R44 que les catégories de temps et de décalage apparaissent toujours ensemble puisque cette contrainte se trouve déjà dans la règle du mode **indicatif**, R36 (p. 308). En effet, ces deux catégories apparaissent toujours ensemble parce qu'elles sont toutes les deux nécessaires quand un verbe est à l'indicatif (cf. Figure 29, p. 221).

2.4.5 La combinatoire des grammèmes d'accord

Les grammèmes d'accord verbe-sujet en nombre et en personne ont une combinatoire simple: ils exigent la présence du grammème **fini**, ce que nous modélisons par les règles de la Figure 103 ci-dessous.

Le participe passé peut s'accorder en genre et en nombre avec son gouverneur (si c'est un nom), le sujet de la copule qui lui est rattachée, ou encore l'objet quand l'auxiliaire AVOIR est utilisé et que cet objet est placé avant l'auxiliaire. Dans tous

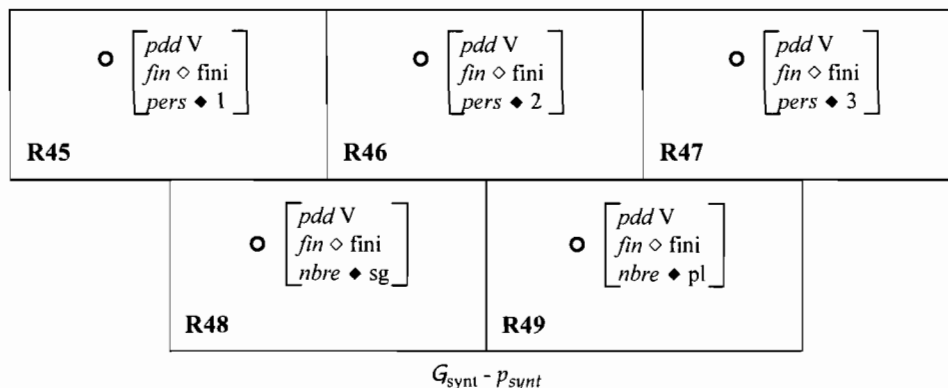


Figure 103 — La combinatoire des grammèmes d'accord (R45 et R49)

ces cas, il se comporte comme un adjectif d'un point de vue morphologique, puisqu'il s'associe aux catégories flexionnelles adjectivales. Pour cette raison, nous laissons de côté ce type d'accord, qui est de peu d'intérêt pour notre étude de l'expression du temps dans la flexion verbale.

2.5 Les règles sagittales

La grammaire de bonne formation syntaxique doit faire l'inventaire des relations syntaxiques qui existent dans une langue donnée et indiquer leur combinatoire (c'est-à-dire, essentiellement, le type de gouverneur dont elles peuvent dépendre, et le type de dépendant qu'elles peuvent avoir). Ce n'est pas notre propos d'étudier la syntaxe du français, ni même celle des verbes de façon exhaustive. Nous renvoyons donc le lecteur à Candito (1999) et Iordanskaja & Mel'čuk (à paraître) pour une discussion en profondeur du problème en français. Sauf avis contraire, les relations syntaxiques que nous utilisons dans notre modèle sont celles de Iordanskaja & Mel'čuk.

2.5.1 Les relations syntaxiques de type *auxiliaire-auxilié*

Les relations syntaxiques qui nous intéressent particulièrement sont celles qui lient les auxiliaires à leur dépendant. On peut reconnaître deux types principaux d'auxiliaires en français: ceux qui imposent le participe passé à leur dépendant (AVOIR_{acc}, ÊTRE_{acc}, AVOIR_{ant} et ÊTRE_{ant}) et ceux qui demandent un verbe à l'infinitif, avec ou sans préposition (ALLER et VENIR). Ces deux groupes se distinguent

notamment par la possibilité ou non de cliticiser le complément de leur dépendant en rattachant ce clitique sur l'auxiliaire :

- (167) a. *Il l'a dit.*
 b. *Ils se le sont dit.*
 c. **Il le va dire.*
 d. **Il le vient de dire.*

La relation syntaxique entre ces deux groupes d'auxiliaires et leur dépendant doit donc être différente. Iordanskaja & Mel'čuk (à paraître) ne se sont intéressés qu'aux relations syntaxiques contrôlées par la valence active des verbes; ils n'ont donc pas touché la question des auxiliaires. Pour le premier groupe, nous stipulerons une relation «Auxiliaire» (Aux). Cette relation ne peut avoir comme gouverneur qu'un verbe, et n'accepte comme dépendant qu'un participe passé :

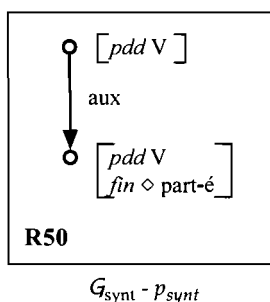


Figure 104 — La relation syntaxique «Aux» (R50)

Les auxiliaires VENIR et ALLER gouvernent tous les deux un verbe à l'infinitif. Cependant, le premier exige la préposition DE, mais pas le second. Bien que Iordanskaja & Mel'čuk ne traitent pas dans leur article des auxiliaires, ils mentionnent néanmoins d'autres constructions où un verbe infinitif, avec ou sans préposition, dépend d'un autre verbe [*devoir partir, commencer à écrire, etc.*]. Ils distinguent deux relations syntaxiques, selon que l'infinitif est gouverné par une préposition ou non: «objet infinitif oblique» (OIO) dans le premier cas et «objet infinitif direct» (OID) dans le second²⁰. Ces deux relations ne se distingueraient que par deux propriétés:

20. Notre traduction de «*oblique-infinitival-objectival*» et «*direct-infinitival-objectival*».

- 1) Le OID est sujet à la cliticisation [*Je peux jouer* → *Je le peux*]; pas le OIO [*Elle commence à s'énerver* → **Elle le commence*]. Les auteurs notent cependant des contre-exemples dans les deux cas [*Je sais nager* → **Je le sais*; *Il a réussi à me convaincre* → *Il y a réussi*].
- 2) Le OID peut être disloqué à gauche [*Partir, je peux*]; pas le OIO [**Préparer le repas, il se dépêche*]. Encore une fois, les auteurs notent qu'il y a des contre-exemples pour les deux [**Partir, je dois*²¹; *Faire un bon repas, Alain réussit toujours*].

Dans le cas de VENIR et ALLER, la seule propriété syntaxique qui les distingue est la capacité du premier à apparaître sans son dépendant :

- (168) a. *Avez-vous parlé à Mlle Xiao ? — Je viens tout juste.*
 b. *Avez-vous parlé à Mlle Xiao ? — * Je vais tout de suite.*

Les deux résistent à la dislocation :

- (169) a. * *De parler à Mlle Xiao, je viens.*
 b. * *Parler à Mlle Xiao, je vais.*

Étant donné que les propriétés distinctives des relations OID et OIO de Iordanskaja & Mel'čuk ne permettent pas une partition absolument nette d'une part, et que d'autre part, une de ces deux propriétés ne permet justement pas de distinguer la relation syntaxique entre VENIR et son dépendant de celle entre ALLER et son dépendant, nous n'utiliserons dans notre modèle qu'une seule relation syntaxique pour les deux cas. Nous appelons cette relation «Infinitif» (Inf). Elle ne peut dépendre que d'un verbe, et son dépendant doit être un verbe à l'infinitif, avec ou sans préposition, ce que nous modélisons par les règles R51 et R52 ci-dessous (page suivante).

2.5.2 *La relation syntaxique copule-attribut*

Nous avons affirmé plus haut que le verbe ÊTRE des constructions résultatives et passives n'est en fait rien d'autre qu'une simple copule (cf. Ch. VI, § 1.2.2, p. 124

21. Iordanskaja & Mel'čuk marquent cet énoncé d'un point d'interrogation. Nous le considérons carrément agrammatical.

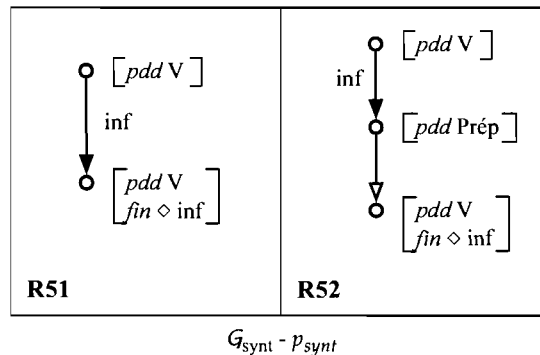


Figure 105 — La relation syntaxique «Inf» (R51 et R52)

et Ch. VI, § 2, p. 132). Par conséquent, la relation qui la lie à son dépendant doit être celle normalement gouvernée par une copule. Il s'agit de la relation «complétive copulo-attributive»²² que Iordanskaja & Mel'čuk utilisent pour les constructions comme *Il est gentil*, *Il est médecin*, *Il est sans argent*, etc. Les auteurs donnent onze propriétés syntaxiques qui caractérisent cette relation. Certaines de ces propriétés ne sont pas pertinentes dans le cas présent (par exemple, elles ne concernent que les dépendants nominaux, ou encore elles font référence à la structure syntaxique profonde, que nous n'avons pas en GUST). Par contre, toutes celles qui sont pertinentes correspondent justement aux propriétés de la relation en jeu dans *Ma voiture est réparée* (qui peut exprimer autant un résultatif qu'un passif):

- 1) Le dépendant de la copule doit s'accorder avec le sujet [*Elle est réparée*].
- 2) Le dépendant peut être cliticisé [*Elle l'est*].
- 3) Cette dépendance n'est pas forcément présente dans toutes les propositions.
- 4) Le gouverneur n'est pas forcément un verbe fini [*Elle a intérêt à être réparée*].
- 5) Le dépendant suit son gouverneur dans la phrase [*Elle est réparée*].
- 6) Le dépendant ne peut pas être disloqué à gauche [**Réparée, elle est*].

Abeillé & Godard (2002) ont aussi démontré de façon systématique que les relations syntaxiques en jeu dans la construction passive et dans les constructions

22. Notre traduction de «copular-attributive-completive».

copule-attribut présentent exactement les mêmes propriétés, et en concluaient qu'il s'agit de la même relation. Nous sommes parfaitement d'accord avec elles. Nous appellerons cette relation «Copule» (Cop); il s'agit en fait de la relation «complétive copulo-attributive» de Iordanskaja & Mel'čuk, que nous avons simplement rebaptisée. Cette relation ne peut dépendre que d'un verbe. Par contre, son dépendant est assez peu contraint. Iordanskaja & Mel'čuk relèvent huit types de dépendants pour cette relation: un adjectif, un nom, une préposition, une conjonction, un nombre, un adverbe, un infinitif ou une proposition (introduite par QUE). Nous constatons qu'il faut ajouter à cette liste: un participe passé. Nous ne présenterons pas toutes les règles sagittales qui décrivent l'éventail des types de dépendants pour cette relation. Les constructions qui nous intéressent dans cette thèse sont celles où le dépendant de la copule est un participe passé (pour le passif et le résultatif) ou une préposition (pour les locutions «EN TRAIN» et «SUR LE POINT»).

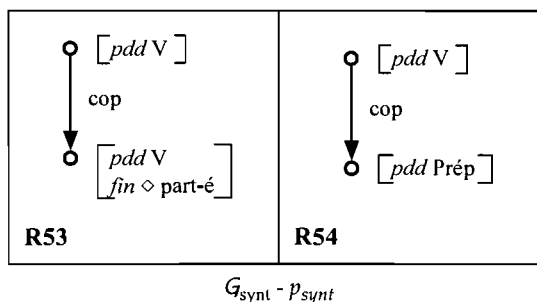


Figure 106 — La relation syntaxique «Cop» (R53 et R54)

2.5.3 Les relations syntaxiques en dehors du noyau verbal

Nous n'avons décrit jusqu'ici que les relations en jeu dans le noyau formé par un auxiliaire et son dépendant ou par la copule et son dépendant. Il y a bien entendu beaucoup d'autres relations syntaxiques dans la langue, et pour chacune il faut décrire les types de gouverneurs et de dépendants avec lesquels elles sont compatibles. Ce travail dépasse largement le cadre de notre thèse. Les quelques règles sagittales ci-dessus suffisent presque pour notre étude. Il nous faut encore décrire deux autres relations importantes pour la flexion verbale en français.

La première est la relation «Préposition» (Prép). Iordanskaja & Mel’čuk ne parlent pas de cette relation, puisqu’ils ne se sont intéressés qu’à la valence des verbes. Il s’agit de la relation qui lie une préposition à son complément [*Il vit à Barcelone, C’est le chapeau de Marie*]. Cette relation est nécessaire notamment pour le gérondif (cf. § 2.4.1, p. 307), pour le complément de VENIR, ainsi que pour les locutions prépositionnelles qui marquent la phase aspectuelle (cf. Figure 94, p. 304). Elle dépend toujours d’une préposition, et son dépendant peut être, pour les cas qui nous intéressent, un verbe non-fini (infinitif ou participe présent) ou un nom.

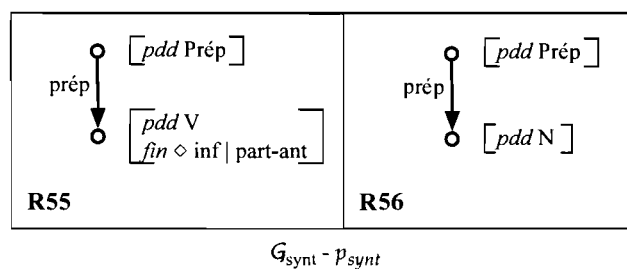


Figure 107 — La relation syntaxique «Prép» (R55 et R56)

Enfin, nous aurons besoin pour notre exposé de la relation «Sujet» (Suj)²³. En effet, le sujet, contrairement aux autres compléments du verbe, doit absolument monter dans la structure syntaxique pour se rattacher à l’auxiliaire, s’il y a lieu (ce phénomène se décrit en $I_{sém-synt}$; nous y reviendrons plus loin, § 3.3, p. 327). La relation «Sujet» ne peut dépendre que d’un verbe fini. Son dépendant typique est un nom (il existe d’autres types de sujet, mais nous nous contenterons du cas typique — voir Iordanskaja & Mel’čuk, à paraître, pour plus de détails). Nous décrivons cette relation en R57 ci-dessous (page suivante).

3 La grammaire d’interface sémantique-syntaxe

La grammaire d’interface sémantique-syntaxe, $I_{sém-synt}$, décrit les liens sémiotiques entre les sens de la langue et leur expression en syntaxe. C’est dans ce

23. Il s’agit à la fois des relations «subjectivale» et «quasi-subjectivale» de Iordanskaja & Mel’čuk, que nous ne distinguons pas dans cette thèse.

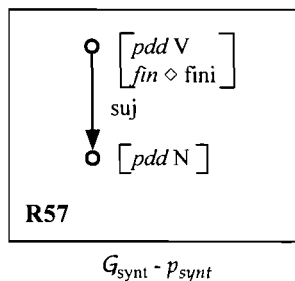


Figure 108 — La relation syntaxique «Suj» (R57)

module que les sémantèmes construits par $G_{\text{sém}}$ (§ 1, p. 283) sont mis en correspondance avec les grammèmes superficiels et les auxiliaires décrits par G_{synt} (§ 2, p. 298). Ainsi, chaque règle de ce module décrit un signe profond (cf. Kahane 2002). C'est donc ici que nous allons modéliser les grammies²⁴ associées aux grammèmes verbaux du français.

Nous avons vu plus haut (§ 1, p. 283) que le signifié des signes linguistiques n'est pas forcément représentable en termes de sémantèmes et de relations entre eux. Nous n'allons considérer ici que les signes profonds dont le signifié peut se représenter par une configuration de sémantèmes, c'est-à-dire par une structure sémantique telle que définie au Chapitre VII (§ 5.1.1, p. 239). Nous laissons aussi de côté les signes dont le signifiant n'est pas syntaxique (en l'occurrence, le présent d'injonction, dont le signifiant est un prosodème). Par conséquent, nous n'allons pas décrire toutes les grammies liées aux grammèmes verbaux.

Dans notre présentation de $G_{\text{sém}}$ (§ 1, p. 283), nous avons regroupé les sens des grammies en trois familles: les sens temporels, les sens phasiques / aspectuels et les sens «modaux». Nous allons regrouper les signes profonds de $I_{\text{sém-synt}}$ de la même façon. Toutefois, il faudra aussi considérer les signes asémantiques. Nous commencerons d'ailleurs notre exposé par ces derniers, mais nous allons d'abord résumer brièvement les caractéristiques formelles de $I_{\text{sém-synt}}$.

24. Rappelons qu'une grammie est une acception particulière d'un grammème (cf. Ch. IV, § 1, p. 46).

3.1 Les caractéristiques formelles de $I_{\text{sém-synt}}$

Nous résumons dans les paragraphes qui suivent les principales caractéristiques formelles de la grammaire d'interface sémantique-syntaxe. Pour une présentation plus complète, le lecteur pourra consulter la section correspondante du chapitre précédent (Ch. VII, § 5.3.2, p. 256).

Le rôle de $I_{\text{sém-synt}}$ est de mettre en correspondance les graphes sémantiques construits par $G_{\text{sém}}$ avec les arbres syntaxiques construits par G_{synt} . Cela se fait par l'ajout de liens entre les nœuds de $G_{\text{sém}}$ et de G_{synt} ainsi que par la saturation de la polarité $p_{\text{sém-synt}}$ laissée en blanc par ces modules. C'est dans ce module que se décrivent les signes profonds à proprement parler, ainsi que la diathèse et le régime des unités lexicales.

Tous les types d'objets qui existent dans $G_{\text{sém}}$ et dans G_{synt} existent aussi dans $I_{\text{sém-synt}}$: nœuds sémantiques (qui représentent les sémantèmes), arcs sémantiques (qui représentent les relations sémantiques), nœuds syntaxiques (qui représentent les lexèmes), arcs syntaxiques (qui représentent les relations syntaxiques) et objets grammémiques (qui représentent les grammèmes superficiels). En plus de ces types d'objets, $I_{\text{sém-synt}}$ comprend également des arcs de correspondance, qui représentent les liens sémiotiques entre les nœuds sémantiques et les nœuds syntaxiques. Tous les objets de $I_{\text{sém-synt}}$ doivent porter une polarité $p_{\text{sém-synt}}$. Par ailleurs, tous les points d'articulation avec $G_{\text{sém}}$ (les nœuds et les arcs sémantiques) portent la polarité d'articulation $p_{\text{sém}}$. De la même façon, les points d'articulation avec G_{synt} (les nœuds et les arcs syntaxiques ainsi que les objets grammémiques) portent la polarité d'articulation p_{synt} . Les polarités d'articulation $p_{\text{sém}}$ et p_{synt} sont toujours blanches dans $I_{\text{sém-synt}}$, c'est pourquoi nous les omettons dans les figures.

Les liens de correspondance apparaissent comme des lignes pointillées reliant les nœuds sémantiques et syntaxiques. Un triangle superposé à cette ligne indique la polarité de l'arc de correspondance. Un même nœud peut être mis en correspondance avec plus d'un nœud du niveau adjacent.

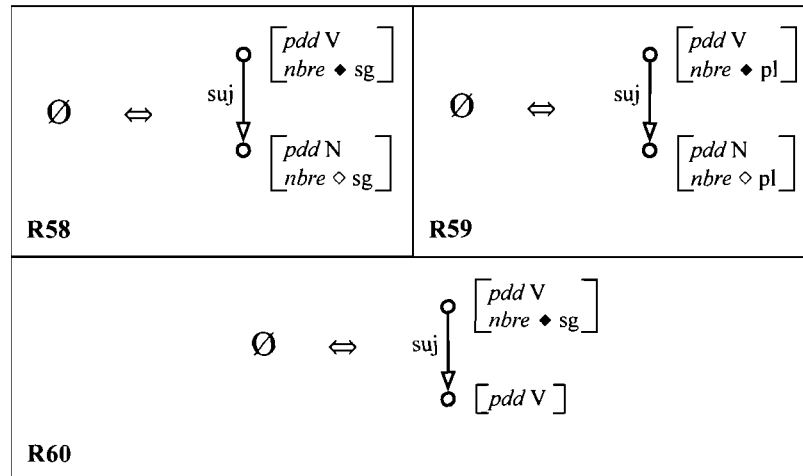
3.2 Les signes asémantiques

Nous entendons par **signe asémantique** un signe dont le signifié n'est pas représentable par des sémantèmes. Cela ne signifie pas que les signes asémantiques ont un signifié vide; seulement, leur signifié est d'une nature qui ne peut être représentée par des sémantèmes. La flexion verbale en français fait intervenir plusieurs types de signes asémantiques. Il y a bien sûr les grammèmes d'accord, qui n'ont jamais aucune autre fonction que de signaler des liens syntaxiques. D'autres grammèmes, en revanche, peuvent avoir des acceptions asémantiques et des acceptions sémantiquement pleines. C'est le cas par exemple des grammèmes de finitude et de mode. Leurs acceptions pleines seront décrites plus loin, dans les sections correspondantes. Leurs acceptions asémantiques sont de deux types: les signes imposés par la structure syntaxique ou le régime des lexies, et les signes liés à la structure communicative. Enfin, certaines lexies sont asémantiques. C'est le cas de la copule ÊTRE_{cop} et des prépositions ou conjonctions régies.

3.2.1 Les grammèmes d'accord

Dans la phrase simple, le verbe doit s'accorder en personne et en nombre avec son sujet. Il s'accorde également en genre et en nombre avec son complément d'objet direct dans certains contextes. Par ailleurs, au participe passé avec ou sans copule, il s'accorde comme un adjectif. Il s'agit là des cas typiques d'accord verbal en français, mais en fait le problème est plus complexe qu'il n'y paraît au premier abord. Nous laissons de côté les cas plus délicats, par exemple les phrases à coordination [*Ma sœur et moi partons en voyage*] ou les sujets en *qui* [*C'est nous qui partons en voyage*], puisque le problème de l'accord est peu pertinent pour notre étude de l'expression du temps grammatical en français.

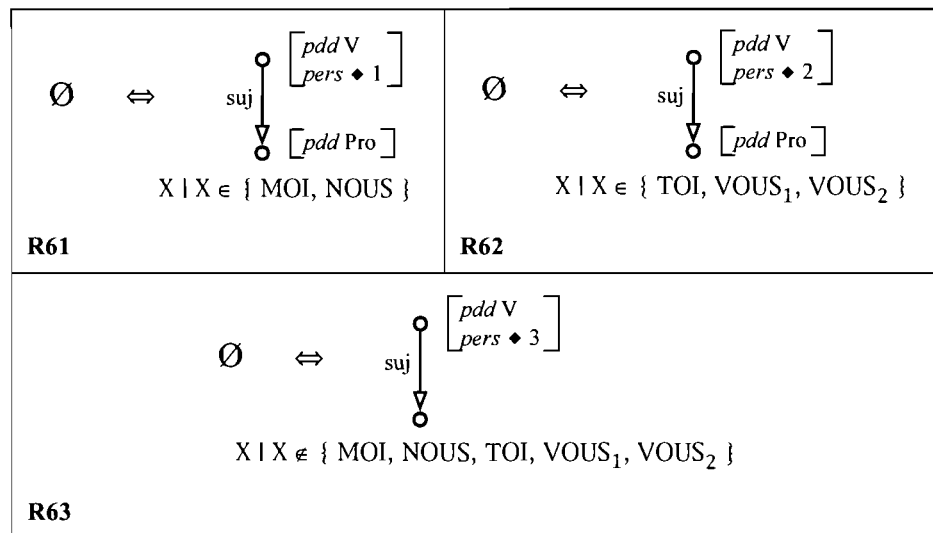
Pour le nombre, l'accord se fait en recopiant sur le verbe l'information flexionnelle portée par le sujet lorsque ce dernier est un nom. Quand le sujet est un autre verbe, alors il impose l'accord au singulier, peu importe sa flexion. C'est ce que modélisent les règles R58 à R60 ci-dessous.



$I_{\text{sém-synt}} - P_{\text{sém-synt}}$

Figure 109 — L'accord verbe-sujet en nombre (R58 à R60)

Pour l'accord personnel, seuls les pronoms MOI, NOUS, TOI, VOUS₁ (pluriel) et VOUS₂ (de politesse) causent un accord à la première ou deuxième personne; toutes les autres lexies déclenchent un accord à la troisième personne. Les règles R61 à R63 ci-dessous modélisent cet accord.



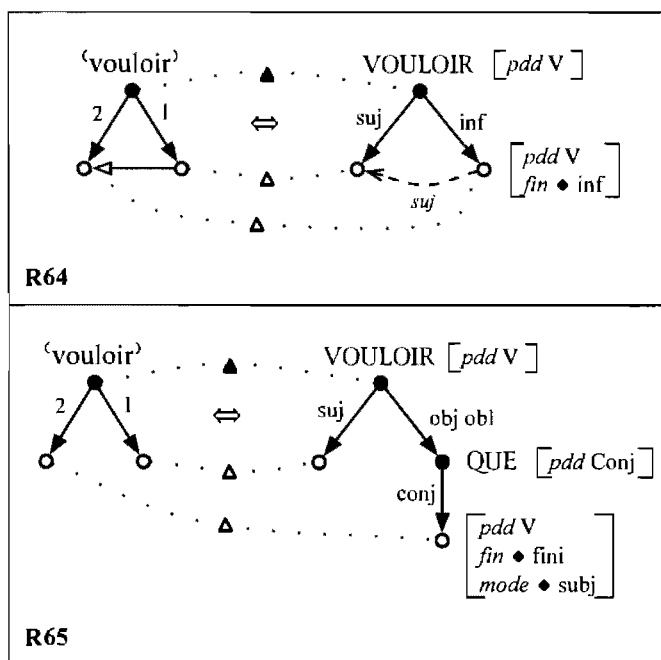
$I_{\text{sém-synt}} - P_{\text{sém-synt}}$

Figure 110 — L'accord verbe-sujet en personne (R61 à R63)

Comme nous l'avons dit plus haut (§ 2.4.5, p. 310), nous ne traitons pas l'accord du participe passé dans cette thèse, puisqu'il présente peu d'intérêt pour notre étude du temps grammatical.

3.2.2 Les signes régis

Une lexie peut imposer à son actant syntaxique un grammème, ou encore forcer l'introduction d'une préposition ou d'une conjonction. Les grammèmes de finitude **fini** et **infinitif**, de même que les grammèmes de mode **indicatif** et **subjonctif**, peuvent être imposés à leur hôte de cette façon. Ils peuvent donc apparaître dans des règles lexicales. Par exemple, les règles R64 et R65 ci-dessous décrivent deux régimes possibles pour la lexie VOULOIR, qui impose à son dépendant soit l'infinitif [*Je veux visiter la Suède*], soit le subjonctif [*Elle veut que j'aïlle en Suède avec elle*]. La règle R65 illustre également comment les prépositions et conjonctions non porteuses de sens, comme QUE, sont introduites. Le mode **indicatif** est aussi



*I*_{sém-synt} - *P*_{sém-synt}

Figure 111 — Le régime de VOULOIR (R64 et R65)

imposé par le régime des lexies. Par exemple, la règle R66 ci-dessous montre un des régimes possibles pour PENSER [*Je pense que la Suède est un beau pays*]. Enfin, quand un verbe est le sommet syntaxique de la phrase, alors il est marqué des grammèmes **fini** et **indicatif**, ce qui est modélisé par la règle R67. Ici, nous faisons appel au trait *pos(ition)* de la règle R17 (p. 301) de $G_{sém}$.

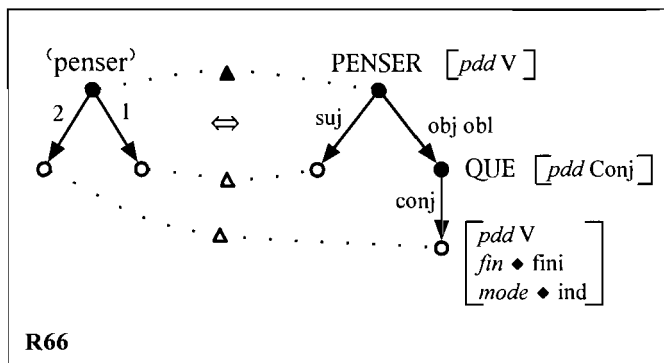


Figure 112 — Le régime de PENSER (R66)

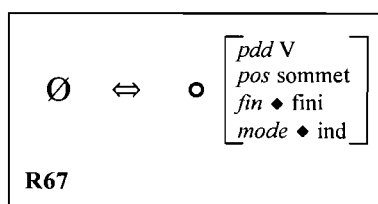


Figure 113 — Le sommet syntaxique (R67)

3.2.3 Les grammies liées à la structure communicative

Les moyens d'expression de la structure communicative sont très variés (ordre, prosodie, etc.)²⁵. Nous n'avons pas étudié la représentation de cette structure en GUST; nous ne pouvons donc pas traiter du problème en profondeur ici. Nous nous intéresserons uniquement aux grammèmes **participe-ant** et **participe-é** en tant que marqueurs de la structure communicative.

Nous avons vu au Chapitre VI (§ 2, p. 132) que les deux participes peuvent apparaître dans le même type de construction: ils dépendent d'un nom, qu'ils modifient à la manière d'un adjectif. Nous avons donné les exemples (25) et (26), que nous reprenons ici en (170) et (171):

(170) *Les journalistes, bombardant le ministre de questions, lui font perdre son sang-froid.*

(171) *Le ministre, bombardé de questions par les journalistes, perd son sang-froid.*

25. Voir Mel'čuk (2001c) pour une revue exhaustive de ces moyens d'expression.

Ce qui unit ces deux participes, c'est le fait qu'ils opèrent comme des modifieurs d'un nom qui se trouve être un de leurs actants sémantiques. Ce qui les distingue, c'est leur diathèse: le premier est actif, le second est passif. Kahane (2002: p. 37) représentait le participe passé, en tant que signifiant du passif, par la règle suivante²⁶:

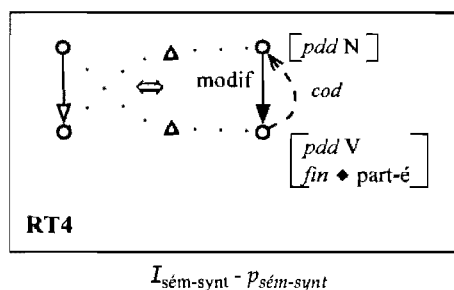


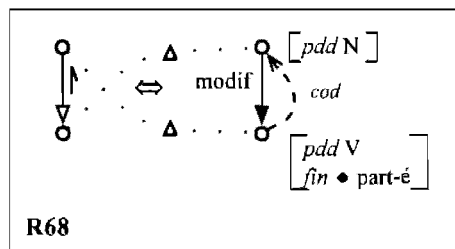
Figure 114 — Le passif selon Kahane (RT4)

Cette règle modélise la promotion de l'objet direct d'un verbe au passif en tant que gouverneur de ce verbe [*la crise a touché des enfants → les enfants touchés (par la crise)*]. Il manque cependant l'information communicative. Il nous semble que la raison pour laquelle un verbe se retrouve comme dépendant de son actant est qu'il est communicativement dépendant de cet actant²⁷. Nous ne savons pas comment bien représenter en GUST cette notion. À défaut d'une meilleure formalisation, nous indiquons cette dépendance par un arc « \vdash » quand elle ne suit pas le même sens que la dépendance sémantique. Nous proposons donc de modifier la règle de Kahane pour obtenir la règle R68 ci-dessous.

En suivant la même logique, on peut représenter le participe présent en opérant la promotion du sujet plutôt que de l'objet [*des enfants se trouvent à l'intérieur de l'immeuble → les enfants se trouvant à l'intérieur de l'immeuble*], ce qu'exprime la règle R69 ci-dessous.

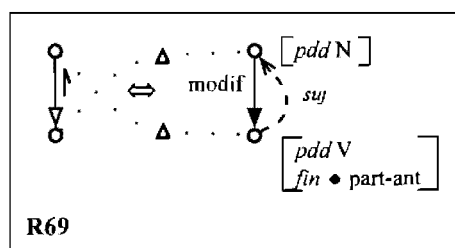
26. Nous faisons un peu entorse à sa règle, qui est antérieure au formalisme GUP de Kahane (2004). La version originale traite aussi la démotion du sujet en agent oblique [*Charles mange la pomme → La pomme est mangée par Charles*].

27. La notion de dépendance communicative a été introduite par Polguère (1990) et reprise par Mel'čuk (2001c).



$$I_{\text{sém-synt}} - p_{\text{sém-synt}}$$

Figure 115 — Le participe passé à valeur de passif (R68)



$$I_{\text{sém-synt}} - p_{\text{sém-synt}}$$

Figure 116 — Le participe présent (R69)

Les règles R68 et R69 modélisent le fait que les participes dépendent du nom qui devrait normalement être leur sujet ou leur objet direct. Les relations «suj» et «cod» sont construites par des règles lexicales comme celles qu'on a vues plus haut pour VOULOIR et PENSER (R64 à R66). Ces relations doivent être effacées pour que l'arbre soit bien formé. Dans nos règles, les arcs syntaxiques en pointillés représentent ce que Kahane (2002: p. 37) appelait une **quasi-dépendance** et qu'il définissait ainsi: «Une quasi-dépendance peut seulement s'unifier avec une dépendance de même étiquette et le résultat de l'unification des deux liens est effacé». Dans le formalisme des grammaires d'unification polarisées, cela se représente par la saturation de la polarité d'interface. Les quasi-dépendances dans $I_{\text{sém-synt}}$ portent une polarité $p_{\text{sém-synt}}$ blanche (ce qui leur permet de s'unifier à une dépendance de même étiquette) mais une polarité p_{synt} noire (ce qui les rend neutres pour G_{synt} , comme si elles étaient «effacées»)²⁸. Nous verrons une illustration de ce mécanisme un peu plus loin, Figure 118a–c, p. 326.

28. Ce n'est pas un hasard si nous représentons les arcs des décompositions sémantiques de $G_{\text{sém}}$ de la même façon (cf. Ch. VII, § 6, p. 267). Ils portent une polarité de construction blanche et une polarité d'interface saturée; ce sont des quasi-dépendances sémantiques.

3.2.4 La copule

La règle du participe passé à valeur de passif ci-dessus (R68) ne modélise que la construction où le passif est utilisé sans copule (nous verrons plus loin, § 3.3, p. 327, que nous modélisons de façon similaire certains des signes de phase aspectuelle). Le passif n'est toutefois pas toujours sans copule; en fait, il est plus fréquent de le trouver avec ÊTRE_{cop}:

(172) *On s'attend à ce que le ministre soit bombardé de questions par les journalistes demain matin.*

Kahane (2002: p. 37) modélisait la copule par la règle ci-dessous²⁹:

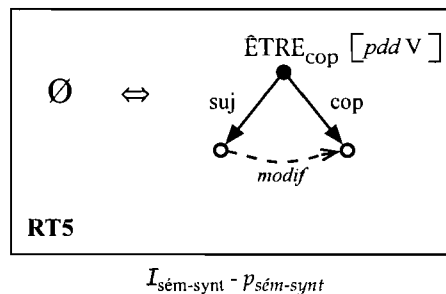


Figure 117 — La copule selon Kahane (RT5)

Cette règle donne la structure syntaxique voulue pour les participes passés à valeur passive ou résultative [*le ministre est bombardé de questions*; *son travail est fini*], les adjectifs [*le chien est fatigué*] ou les prépositions pleines [*le disque est sur la table*]. Cependant, elle n'indique pas d'où doivent venir les informations de temps grammatical que peut porter cette copule, ni comment elle doit se rattacher à un éventuel gouverneur. Considérons l'exemple suivant:

(173) *Je crois qu'elle est aimée de tous.*

Ici, la construction passive avec copule se trouve dans une subordonnée. La copule porte également les grammèmes **simultané** et **non-décalé**. La représentation sémantique de cette phrase, sans l'information grammaticale de temps, est celle qu'on peut voir à la Figure 118a ci-dessous. Des règles de lexicalisation pour 'aimer', 'tous' et 'elle', que nous n'avons pas présentées mais qui sont triviales, per-

29. Nous la reformulons en termes de structure polarisée. Nous l'avons également modifiée légèrement; Kahane décrit la combinatoire morphologique des signes dans $I_{sém-synt}$.

mettent de neutraliser une partie de ce réseau et de construire le fragment d'arbre correspondant, comme on peut le voir à la Figure 118b. La règle R68 du passif pro-met le complément d'objet direct en position de gouverneur de AIMER, ce qui donne la structure à la Figure 118c. En appliquant ensuite la règle RT5 de Kahane, on obtient la structure de la Figure 118d.

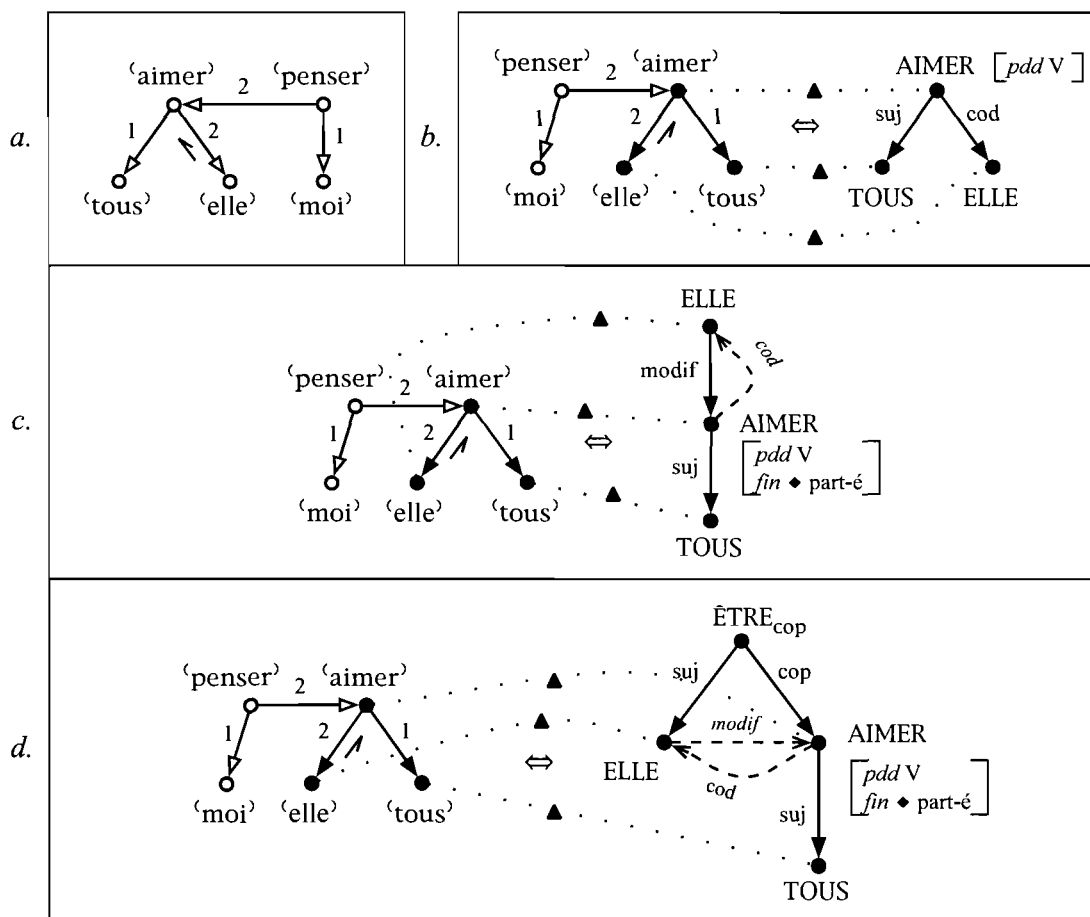
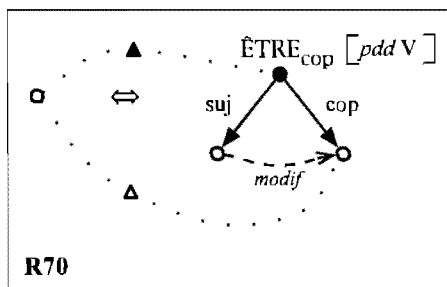


Figure 118 — La dérivation d'un passif avec copule dans $I_{sém-synt}$

Maintenant, si on tente de neutraliser 'penser' avec la règle R66 (p. 322), on n'aura d'autre choix que d'en relier le lexème correspondant à AIMER, et non ÊTRE_{cop}, puisque c'est ce nœud qui est marqué d'un lien de correspondance avec le second actant sémantique de 'penser'. Il en résulte une représentation syntaxique mal formée où AIMER a deux gouverneurs: ÊTRE_{cop} et PENSER. La solution à ce problème est de modifier la règle RT5 pour que la copule soit mise en correspondance avec le sémantème qu'exprime son actant syntaxique:

VIII. Un modèle formel de la conjugaison française



$I_{sém-synt} - P_{sém-synt}$

Figure 119 — La copule dans $I_{sém-synt}$ (R70)

De cette façon, la représentation qu'on obtient est la suivante :

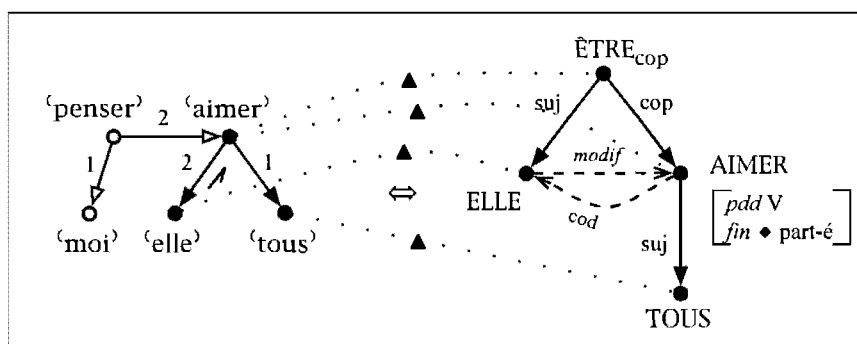


Figure 120 — Une double correspondance pour les constructions à copule

Cette fois, 'aimer' est mis en correspondance non seulement avec AIMER, mais aussi avec ÊTRE_{cop}. Ainsi, il y a deux points d'ancrage possibles pour le complément de PENSER. Un seul des deux (ÊTRE_{cop}) permet la construction d'un arbre syntaxique qui pourra être validé par G_{synt} . Nous allons voir dans les prochaines sections que la copule et les doubles correspondances sont aussi utiles pour modéliser les signes temporels et aspectuels.

3.3 Les signes de phase aspectuelle et le résultatif

Nous avons vu dans notre présentation de $G_{sém}$ qu'il y a dans notre modèle six sémantèmes de type phasique ou aspectuel. Il y a 'aller', 'en train' et 'avoir_{acc}', qui dénotent les périodes avant, pendant et après un fait, 'sur le point' et 'venir', qui dénotent les périodes immédiatement avant ou immédiatement après un fait, et 'résultatif', qui dénote l'état résultant d'un fait (cf. § 1.4, p. 289).

Les auxiliaires et locutions prépositionnelles AVOIR_{acc}, ALLER, 「EN TRAIN」, VENIR et 「SUR LE POINT」 expriment les sens phasiques correspondants. Ils ne forment pas une catégorie flexionnelle au sens où nous l'entendons puisqu'ils ne sont pas mutuellement exclusifs [*Ils vont avoir fini avant nous*] et ne sont pas forcément de la même partie du discours que la base verbale [*Je les ai trouvés en train de siroter des mojitos*].

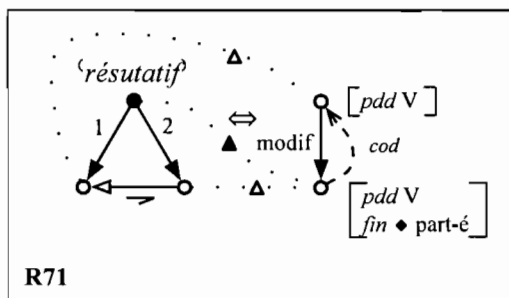
Le participe passé à valeur de résultatif n'est pas exactement un signe de phase aspectuelle, mais il a quand même un sens de nature aspectuelle, c'est pourquoi nous le regroupons avec les marqueurs de phase ici. Comme nous venons de parler du participe passé à valeur passive, nous allons commencer notre exposé des signes aspectuels par le résultatif, qui se traite de façon similaire au passif. Nous passerons ensuite aux locutions prépositionnelles et aux auxiliaires de phase aspectuelle.

3.3.1 *Le résultatif*

Le sens 'résultatif' que nous avons décrit plus haut (§ 1.4, p. 289) [*La voiture est réparée* (= 'elle n'est plus brisée')] s'exprime par un participe passé. La copule n'est pas nécessaire dans ce type de construction [*Les voitures entrent au garage brisées et en ressortent réparées*]; elle ne fait donc pas partie du signifiant de ce signe et elle se combine à ce signe selon la même mécanique que pour le passif. Comme le passif, le résultatif promet son objet direct à la position de gouverneur du verbe, ce que modélise la règle R71 ci-dessous. Cette règle établit un lien de correspondance entre le nœud 'résultatif' et le verbe au participe passé, qui lui-même a déjà un correspondant dans la représentation sémantique. Le but est le même que pour la règle R70 (p. 327): il s'agit d'assurer un ancrage pour les signifiants de prédicats qui auraient comme argument 'résultatif'.

3.3.2 *Les locutions prépositionnelles de phase aspectuelle*

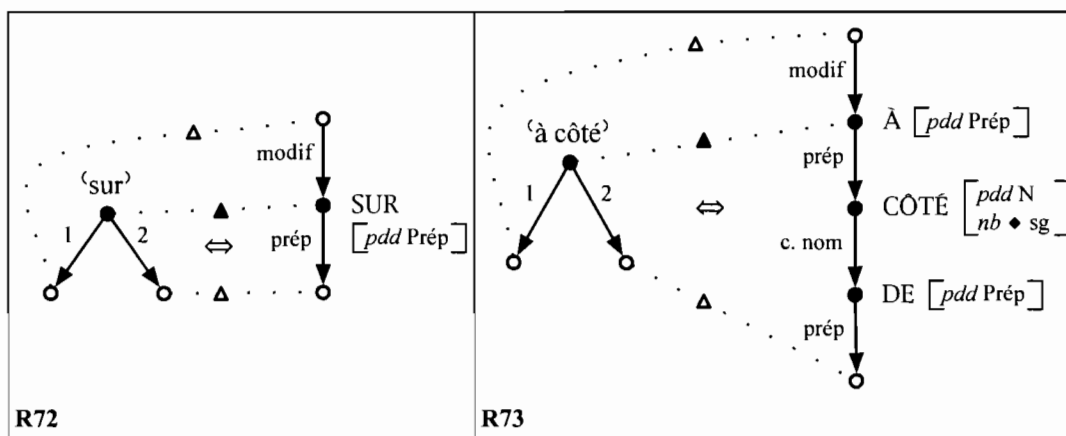
Les sémantèmes 'en train' et 'sur le point' que nous avons décrits plus haut (§ 1.4, p. 289) s'expriment par des locutions prépositionnelles. La copule ne fait pas partie du signifiant de ces signes, puisqu'elle n'est pas nécessaire dans certains



$I_{sém-synt} - P_{sém-synt}$

Figure 121 — Le participe passé à valeur résultative dans $I_{sém-synt}$ (R71)

contextes syntaxiques (cf. Ch. VI, § 5.3, p. 169). Elle ne doit donc pas apparaître dans les règles qui décrivent ces signes. Avant de voir comment se modélisent les locutions prépositionnelles de phase aspectuelle, il est utile de montrer comment se traitent les prépositions pleines en général. Les deux règles de la Figure 122 ci-dessous décrivent les unités lexicales SUR [*Passe-moi le livre sur la table*] et $\overline{\text{À CÔTÉ}}$ [*Passe-moi le livre à côté de toi*]. Leur traitement est le même, si ce n'est que dans le cas de la locution $\overline{\text{À CÔTÉ}}$ il faut décrire toute la structure syntaxique de la locution, y compris les grammèmes qui en font partie (en l'occurrence, le grammème **singulier** sur le nœud CÔTÉ). Dans les deux cas, le premier actant sémantique (dans nos exemples, il s'agit de 'livre') domine syntaxiquement la préposition.



$I_{sém-synt} - P_{sém-synt}$

Figure 122 — Les prépositions $\overline{\text{À CÔTÉ}}$ et SUR dans $I_{sém-synt}$ (R72 et R73)

Les locutions prépositionnelles de phase aspectuelle, contrairement aux autres prépositions pleines du français, n'ont qu'un seul actant sémantique, et cet actant doit dépendre de la locution prépositionnelle, et non en être le gouverneur. C'est

donc le sujet de cet actant qui est promu à la position de gouverneur de la locution [*Pierre lit* ⊕ *en train* → (*J'ai trouvé*) *Pierre en train de lire*]. C'est ce que modélisent les deux règles de la Figure 123 ci-dessous.

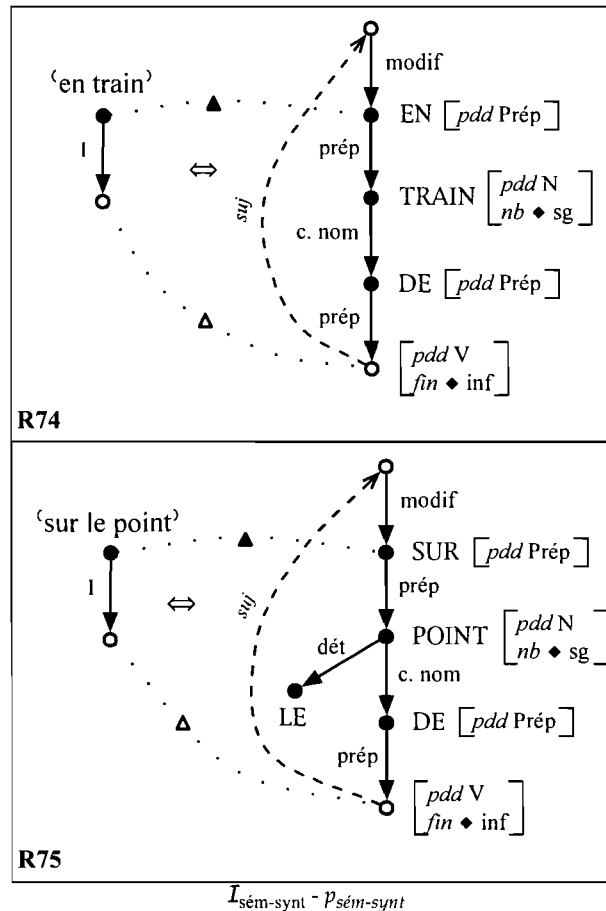


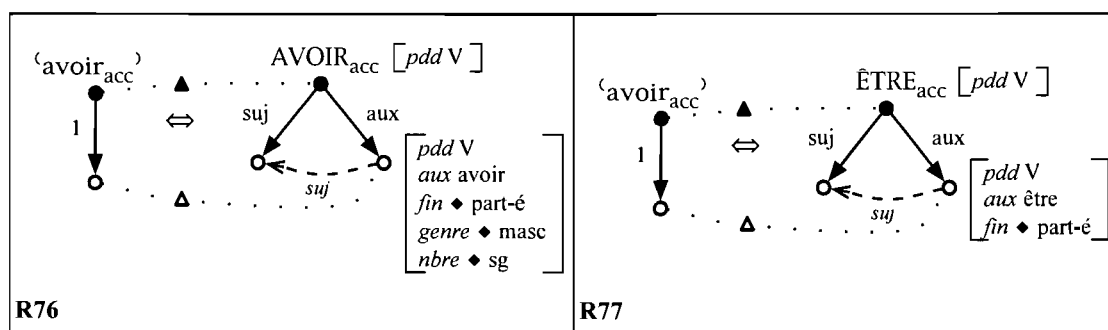
Figure 123 — Les prépositions «EN TRAIN» et «SUR LE POINT» dans $I_{sém-synt}$ (R74 et R75)

Les locutions et prépositions sémantiquement pleines, comme celles présentées ici, si elles doivent être le sommet d'une proposition, ont besoin d'une copule [*Mon livre est à côté du lit*]. Cette copule ne sert qu'à offrir un support verbal à la préposition, qui ne pourrait pas occuper la position de sommet syntaxique; elle ne fait pas partie du signifiant de ces prépositions. C'est la règle R70 (p. 327) qui introduit cette copule pour «EN TRAIN» et «SUR LE POINT» ainsi que pour toutes les autres prépositions pleines comme SUR, DANS, EN-DESSOUS, «À CÔTÉ», etc. Nous

allons voir plus loin (Figure 126, p. 333) une illustration de la montée du sujet du noyau verbal jusqu'à la copule.

3.3.3 Les auxiliaires de phase aspectuelle

Les auxiliaires de phase aspectuelle sont AVOIR_{acc}, ÊTRE_{acc}, ALLER et VENIR. Les sens qu'ils expriment ont été décrits dans $G_{sém}$ (§ 1.4, p. 289). Tous les auxiliaires ont la particularité de faire monter sur eux-même le sujet de leur dépendant [*il mange* → *il vient de manger*]. Kahane (2002: p. 36) modélise ce phénomène par une quasi-dépendance; nous faisons la même chose. L'auxiliaire qui marque l'accompli varie selon le verbe qui en dépend; il faut donc avoir accès à cette information dans la règle. Cet auxiliaire et le verbe qui en contrôle le choix sont toujours directement connectés dans l'arbre syntaxique, peu importe le nombre d'auxiliaires qu'il peut y avoir. Kahane (2002: p. 36) représente cela au moyen d'un opérateur qui retourne la valeur d'un trait; nous préférons éviter ici l'usage de cet opérateur, qui complique le formalisme. Nous avons donc besoin de deux règles pour modéliser la sélection de l'auxiliaire:



$I_{sém-synt} - P_{sém-synt}$

Figure 124 — La phase accomplie dans $I_{sém-synt}$ (R76 et R77)

Le genre et le nombre du participe passé n'est pas spécifié dans la règle R77, puisqu'il doit y avoir accord. Par contre, dans R76 il n'y a pas d'accord³⁰; le genre et le nombre font partie du signifiant de ce signe. Le fait que la finitude, le mode et le temps ne soient pas spécifiés pour l'auxiliaire dans les règles ci-dessus permet

30. Rappelons que nous laissons de côté le problème de l'accord du participe passé dans cette thèse. Lorsque le complément du participe est placé avant AVOIR, il y a accord.

de générer des formes comme *Ils vont avoir mangé*, où l'auxiliaire de la phase accomplie est à l'infinitif, ou *Je ne pense pas qu'ils aient mangé*, où il est au subjonctif. Par contre, son dépendant doit porter le grammème **participe-é**, qui fait partie du signifiant de ces signes.

Le traitement des deux autres auxiliaires de phase est tout à fait similaire, si ce n'est de leur structure syntaxique qui est légèrement différente :

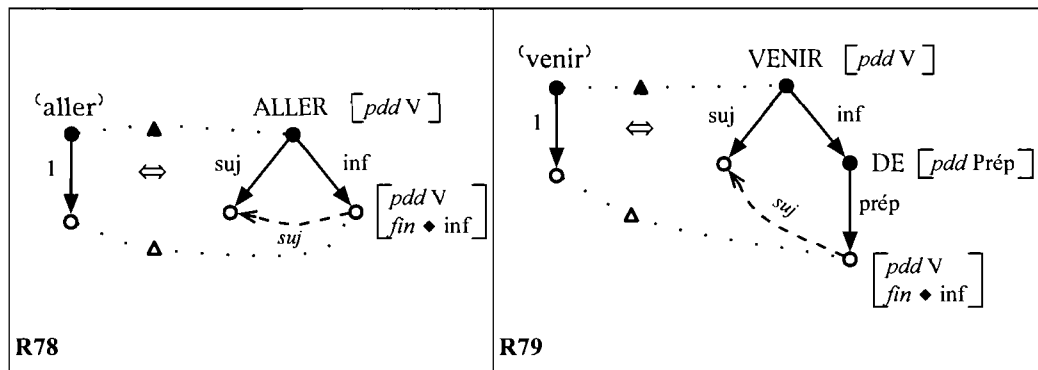


Figure 125 — La phase prospective et le passé récent dans $I_{\text{sém-synt}}$ (R78 et R79)

ALLER et VENIR ne sont pas marqués en finitude, mode et temps, bien qu'ils ne puisse se combinent librement aux grammèmes de ces catégories. Ces contraintes sont déjà décrites en G_{synt} par la règle R25 (p. 305), elles n'ont donc pas à être décrites ici.

Les règles R74 à R79 ci-dessus, qui décrivent les signes de phase aspectuelle, ainsi que R70 (p. 327), qui modélise la copule, suffisent à assurer la montée du sujet, peu importe le nombre d'auxiliaires présents. Nous allons illustrer le phénomène à l'aide de la phrase (174) ci-dessous, où le sujet passe successivement de DORMIR à \ulcorner EN TRAIN \urcorner ³¹, puis à ÊTRE, et finalement à ALLER.

(174) *Ils vont être en train de dormir.*

La représentation sémantique de (174) (Figure 126a)³² est partiellement neutralisée par les règles de lexicalisation de DORMIR et EUX (omises parce que trivialisées).

31. En fait, le sujet de DORMIR ne devient pas le sujet de \ulcorner EN TRAIN \urcorner , mais son gouverneur.

32. Nous ignorons ici le temps grammatical, dont nous parlerons plus loin (§ 3.4, p. 334).

les), ce qui donne la représentation syntaxique partielle de la Figure 126b. L'application de la règle R74 (p. 330) de 'EN TRAIN' a pour effet de promouvoir le sujet de DORMIR en position de gouverneur de la locution prépositionnelle (Figure 126c)³³. Il est impossible à ce stade-ci d'appliquer la règle R78 de ALLER puisque cet auxiliaire demande un complément verbal. Il faut donc introduire une copule en appliquant la règle R70 (p. 327). On obtient alors la structure à la Figure 126d (page suivante). Enfin, on peut appliquer la règle R78 qui introduit l'auxiliaire ALLER, ce qui nous donne la structure de la Figure 126e (page suivante).

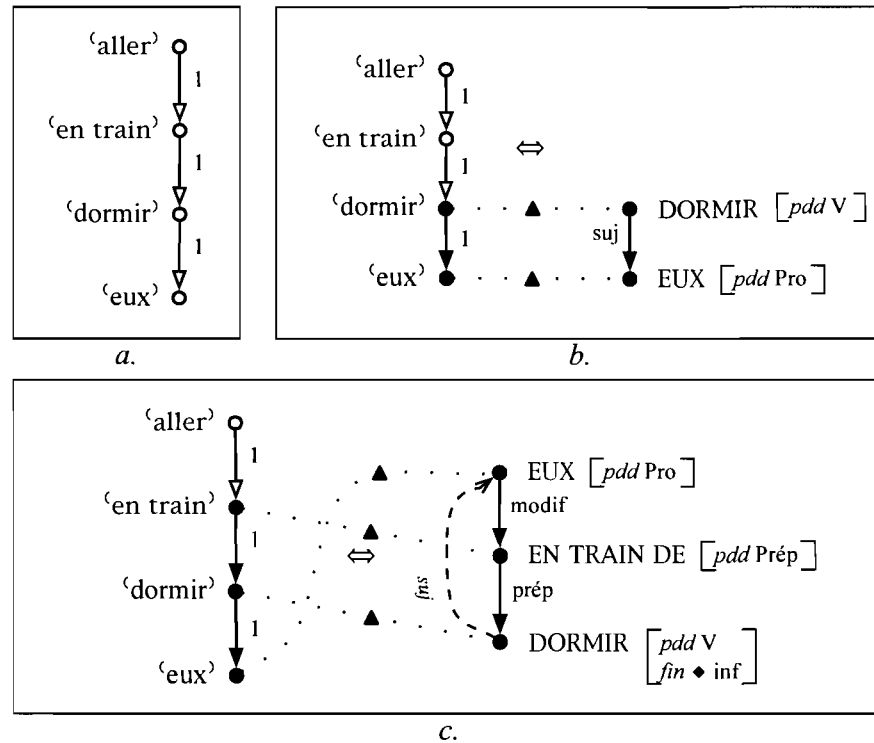


Figure 126 — Les étapes de la dérivation de Ils vont être en train de dormir

33. Nous réduisons *en train de* à un seul nœud pour sauver de l'espace.

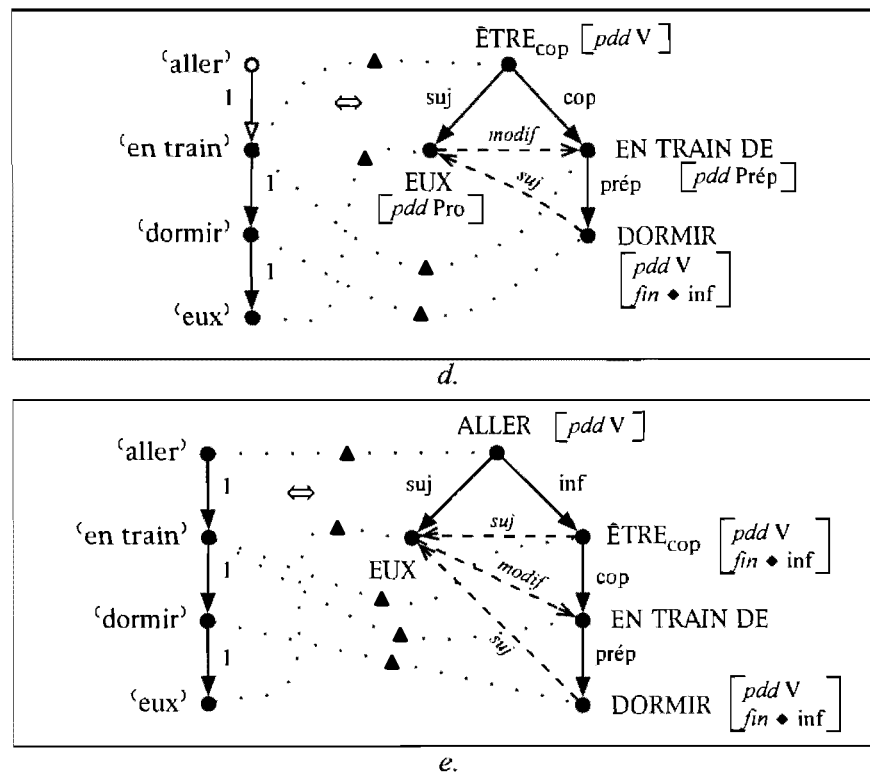


Figure 126 — Les étapes de la dérivation de *Ils vont être en train de dormir*

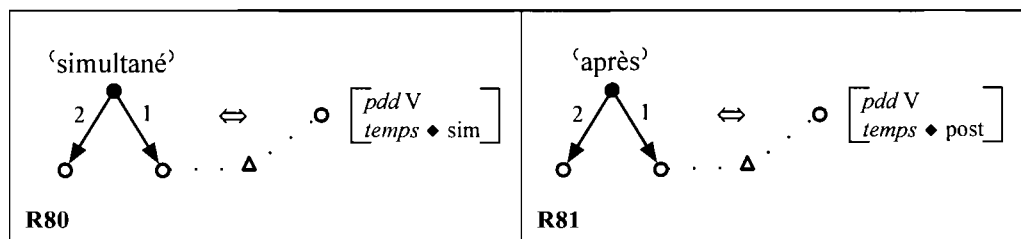
3.4 Les grammies à valeur temporelle

Nous avons vu plus haut (§ 1.3, p. 286) qu'il n'y a que trois sémantèmes temporels en jeu dans la flexion verbale: 'simultané', 'antérieur' et 'postérieur'. Ils correspondent aux trois grammèmes de la catégorie flexionnelle de temps. Selon qu'ils situent les faits par rapport à un repère passé ou non, ils déclenchent aussi l'utilisation des grammèmes **décalé** ou **non-décalé** de la catégorie flexionnelle de décalage (cf. Ch. VI, § 6.4, p. 202 et Ch. VIII, § 1.3, p. 286).

3.4.1 L'acception de base des grammies de temps

Dans leur acception de base, les grammèmes **simultané** et **antérieur** expriment les prédicats 'simultané' et 'postérieur' décrits plus haut (§ 1.3, p. 286), qui situent un fait par rapport au moment de référence. Cela se modélise par les règles R80 et R81 ci-dessous. Puisque le point de repère peut être soit le moment d'énon-

ciation, soit un autre fait, le nœud qui le représente ne porte pas d'étiquette dans cette règle. Il pourra alors s'unifier avec n'importe quel nœud sémantique.



$I_{\text{sém-synt}} - \mathcal{P}_{\text{sém-synt}}$

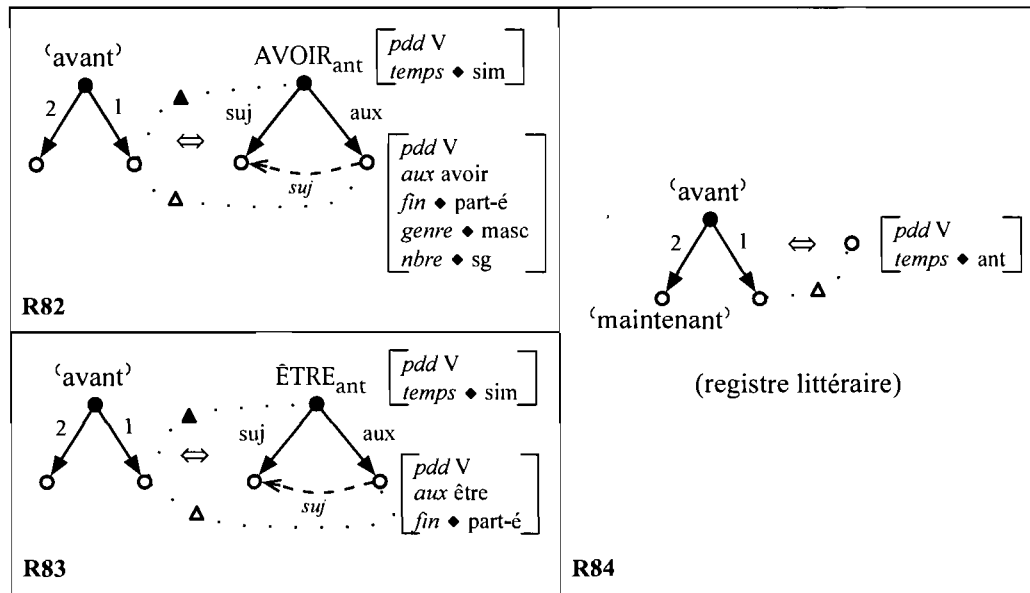
Figure 127 — L'expression des sens 'simultané' et 'postérieur' (R80 et R81)

Le sémantème 'antérieur' peut s'exprimer en français soit par AVOIR_{ant} (ou ÊTRE_{ant}), soit par le passé simple. Il y a donc trois règles concurrentes qui mettent en correspondance le même prédicat avec chacun de ces marqueurs. Les règles R82 et R83 ci-dessous (page suivante) introduisent un auxiliaire (selon le syntactique de l'auxilié) et lui imposent le temps **simultané**. Ils imposent aussi la finitude **participe-é** au verbe qui en dépend. Ces grammèmes superficiels sont vidés de leur sens puisqu'ils font partie du signifiant du grammème profond **antérieur**. Les auxiliaires de temps forcent la montée du sujet, selon la même mécanique que les auxiliaires de phase aspectuelle (cf. § 3.3.3, p. 331). La règle R84 (page suivante), quant à elle, introduit simplement le grammème **antérieur**, qui pourra être mis en correspondance avec le suffixe du passé simple par la grammaire d'interface syntaxe-morphotopologie (mais seulement si l'actant de 'avant' est 'maintenant').

3.4.2 L'acception de base des grammies de décalage

Dans leur acception de base, les grammèmes de décalage n'expriment pas en soi des sémantèmes, mais plutôt des configurations de sémantèmes. Ils ne saturent donc aucun objet du niveau sémantique. Ils construisent néanmoins des objets au niveau syntaxique.

Dans son acception de base, le grammème **non-décalé** indique que le repère temporel est le moment d'énonciation ou un autre moment dans le futur. Si, au contraire, un fait est situé temporellement par rapport à un autre fait qui lui-même



$I_{s\acute{e}m-synt} - P_{s\acute{e}m-synt}$

Figure 128 — L'expression du sens 'antérieur' dans $I_{s\acute{e}m-synt}$ (R82 à R84)

se trouve avant le moment d'énonciation, alors le verbe qui exprime ce premier fait porte le grammème **décalé**. C'est ce que modélisent les règles R85 à R87 ci-dessous³⁴.

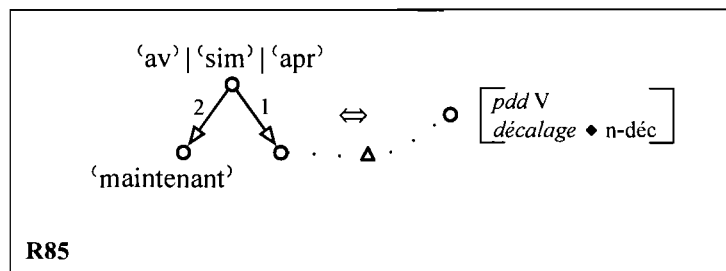
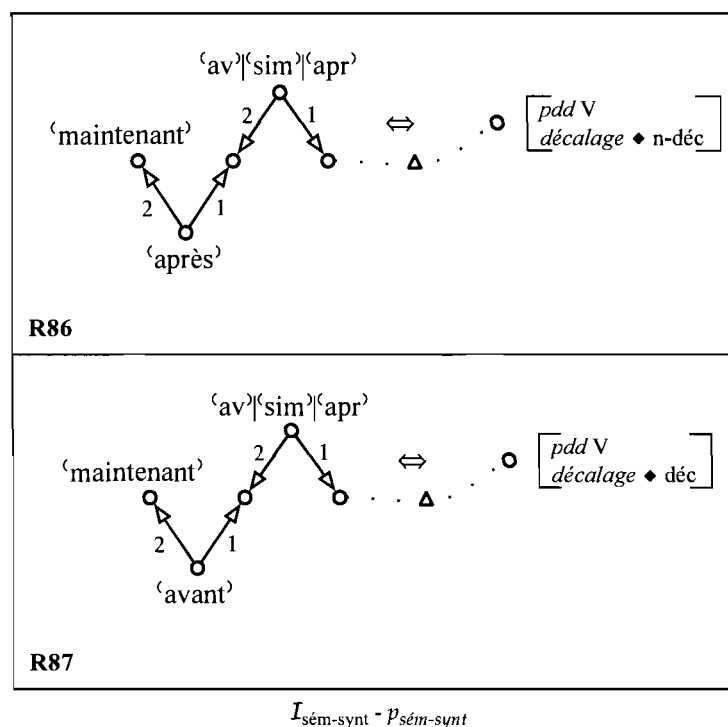


Figure 129 — L'acceptation de base des grammèmes de décalage (R85 à R87)

34. Les abréviations «'av' | 'sim' | 'apr'» signifient «'avant' | 'simultané' | 'après'».



$I_{\text{sém-synt}} - P_{\text{sém-synt}}$

Figure 129 — L'acceptation de base des grammèmes de décalage (R85 à R87)

Nous allons illustrer le fonctionnement des règles temporelles à partir de la phase suivante :

(175) *Ils ont dit à Paul que tu aurais mangé.*

Cette phrase est intéressante parce qu'elle met en jeu plusieurs types de signes profonds. Le premier auxiliaire (*ont*) est une forme de $AVOIR_{\text{ant}}$, qui marque l'antériorité par rapport au moment d'énonciation. Le second (*aurais*) dénote la phase aspectuelle accomplie. La période dénotée par ce signe est située temporellement par rapport à 'dire' : elle lui est postérieure. On pourra ainsi observer la réalisation d'un sens temporel par un auxiliaire (*ont*) et par un grammème superficiel (*aurais*), l'usage de **décalé** (*aurais*) et de **non-décalé** (*ont*), ainsi que la composition d'un signe aspectuel ($AVOIR_{\text{acc}}$) et de signes temporels (**postérieur** et **non-décalé**) dans *aurais*.

La Figure 130³⁵ ci-dessous montre la représentation partielle dans $I_{\text{sém-synt}}$ de la phrase (175) après saturation de la structure sémantique par les règles de lexicalisation de 'toi', 'manger', 'eux', 'Paul' et 'dire' (que nous n'avons pas décrites mais

qui sont triviales), ainsi que l'application des règles R76 (p. 331) pour saturer 'avoir_{acc}', R81 (p. 335) pour 'après' et R82 (p. 336) pour 'avant'. Le sémantème 'maintenant' n'est pas polarisé (ce qui est représenté par la couleur grise). En effet, nous considérons ce sens comme une partie inhérente de toute situation de communication; il est donc toujours présent implicitement et n'a pas besoin d'être construit ou exprimé en syntaxe.

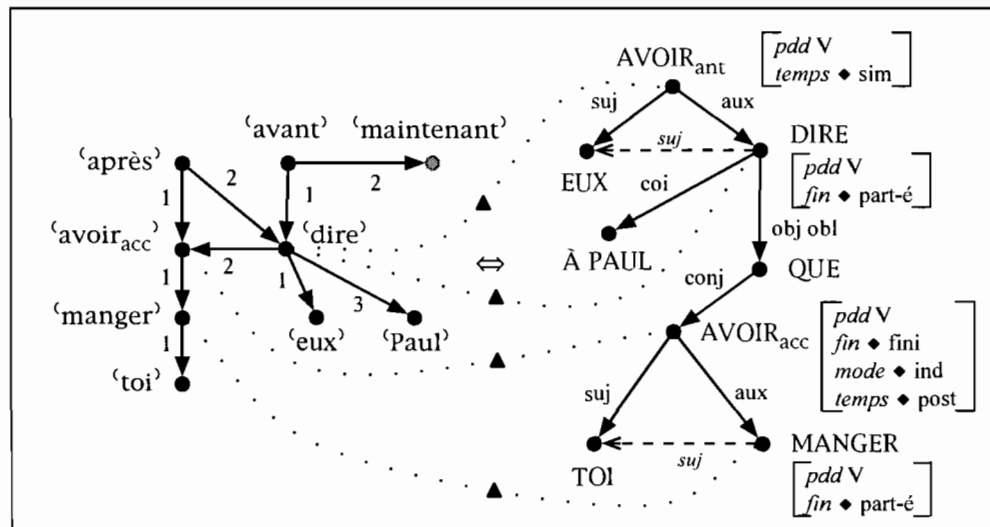


Figure 130 — La représentation partielle de (175) dans $I_{sém-synt}$

Cette structure est complètement saturée. Donc, du point de vue de $I_{sém-synt}$, le travail est terminé. Pourtant, il manque plusieurs grammèmes: la finitude et le mode pour le sommet syntaxique, les grammèmes d'accord pour les deux auxiliaires ainsi que les grammèmes de décalage. Cependant, il ne faut pas oublier que $I_{sém-synt}$ est articulée avec G_{synt} (cf. Ch. VII, § 5.2, p. 251 et Ch. VII, § 5.3.2, p. 256). Donc, toute structure générée par $I_{sém-synt}$ doit être vérifiée par la grammaire de bonne formation syntaxique. Or, ce module exige la présence des grammèmes manquants; par le mécanisme de va-et-vient entre les modules décrit au chapitre précédent (Ch. VII, § 5.3.2, p. 256), G_{synt} peut donc forcer l'application de règles additionnelles dans $I_{sém-synt}$, de la façon suivante:

35. Afin d'alléger la figure, nous omettons les traits des nœuds non verbaux, ainsi que leurs liens de correspondance. De plus, nous réduisons à *Paul* à un seul nœud en syntaxe. Enfin, nous n'indiquons pas le genre et le nombre des participes passés; les deux portent les grammèmes **masculin** et **singulier**.

- La présence de AVOIR_{acc} et AVOIR_{ant} dans la structure syntaxique déclenchera les règles R23 (p. 305) et R29 (p. 306) de $\mathcal{G}_{\text{synt}}$, qui décrivent la combinatoire morphologique de ces auxiliaires et qui demandent une finitude. Celle de AVOIR_{acc} est déjà fournie par le régime de DIRE³⁶. Celle de AVOIR_{ant}, qui n'est gouverné par aucun lexème, ne peut être fournie que par la règle R67 (p. 322) de $\mathcal{I}_{\text{sém-synt}}$, qui impose les grammèmes **fini** et **indicatif** au sommet syntaxique.
- La présence de **fini** sur les auxiliaires aura pour effet, en $\mathcal{G}_{\text{synt}}$, de forcer la présence de grammèmes d'accord, grâce à la règle R32 (p. 307), qui décrit la combinatoire morphologique de **fini**.
- Cette nouvelle exigence de $\mathcal{G}_{\text{synt}}$ déclenchera l'application des règles R58, R59, R62 et R63 (pp. 320–320) de $\mathcal{I}_{\text{sém-synt}}$, qui fourniront les grammèmes d'accord pour AVOIR_{ant} et AVOIR_{acc}.
- Le grammème **indicatif**, que portent AVOIR_{ant} à cause du régime de DIRE et AVOIR_{acc} à cause de la règle R67 (p. 322) qui impose ce mode au sommet syntaxique, déclenchera l'application en $\mathcal{G}_{\text{synt}}$ de R36 (p. 308), qui décrit la combinatoire morphologique de ce grammème. Cela aura pour effet d'exiger la présence de grammèmes de temps et de décalage pour les deux auxiliaires.
- Les auxiliaires portent déjà un grammème de temps. Celui de AVOIR_{ant} a été fourni par la règle R82 (p. 336) de $\mathcal{I}_{\text{sém-synt}}$, qui sature le sémantème 'antérieur' et qui construit l'auxiliaire (cet auxiliaire porte toujours le temps **simultané**, qui fait partie de ce phrasème grammémique). L'auxiliaire AVOIR_{acc}, quant à lui, porte le grammème **postérieur** à cause de la règle R81 (p. 335) qui s'est appliquée pour saturer le sémantème 'postérieur'.
- Puisque 'avoir_{acc}' est situé par rapport à 'dire', qui lui-même est situé dans le passé, la règle R85 (p. 336) de $\mathcal{I}_{\text{sém-synt}}$ pourra s'appliquer pour fournir au nœud correspondant le grammème **décalé**. Par contre, 'dire' est situé directement par rapport à 'maintenant'. C'est donc la règle R87 (p. 337) qui doit s'appliquer pour fournir au nœud correspondant le grammème **non-**

36. Nous n'avons pas décrit ce signe, mais il est similaire à PENSER, décrit en R66 (p. 322).

décalé. Or, 'dire' a deux nœuds correspondants en syntaxe à cause de la règle R82 (p. 336), qui construit l'auxiliaire AVOIR_{ant} et le met en correspondance avec le même sémantème que son auxilié. Il y a donc deux sites d'ancrage possibles pour **non-décalé**: AVOIR_{ant} et DIRE. Du point de vue de $I_{\text{sém-synt}}$, les deux sites sont parfaitement valides. C'est G_{synt} qui force le choix de l'auxiliaire comme site d'ancrage pour ce grammème.

- Si le grammème **non-décalé** était placé sur DIRE, cela déclencherait la règle R44 (p. 310) de G_{synt} , qui décrit la combinatoire de **non-décalé**. Ce grammème exige la présence du mode **indicatif**. Or, la combinatoire de **indicatif**, décrite dans la règle R36 (p. 308) de G_{synt} , demande le grammème de finitude **fini**. Puisque DIRE porte déjà la finitude **participe-é**, il serait impossible de saturer la structure.
- Par contre, si **non-décalé** est placé sur AVOIR_{ant}, alors l'application des règles de G_{synt} R44 (p. 310) et R36 (p. 308), qui décrivent la combinatoire de **non-décalé** et de **indicatif**, se fera sans problème. Donc, bien que 'dire' ait deux correspondants au niveau syntaxique, il n'y a pas de choix pour l'ancrage du grammème de décalage: il ne peut aller que sur AVOIR_{ant}.

Après ce ping-pong grammatical entre les modules $I_{\text{sém-synt}}$ et G_{synt} , on obtient finalement la structure ci-dessous, qui est entièrement saturée et bien formée³⁷.

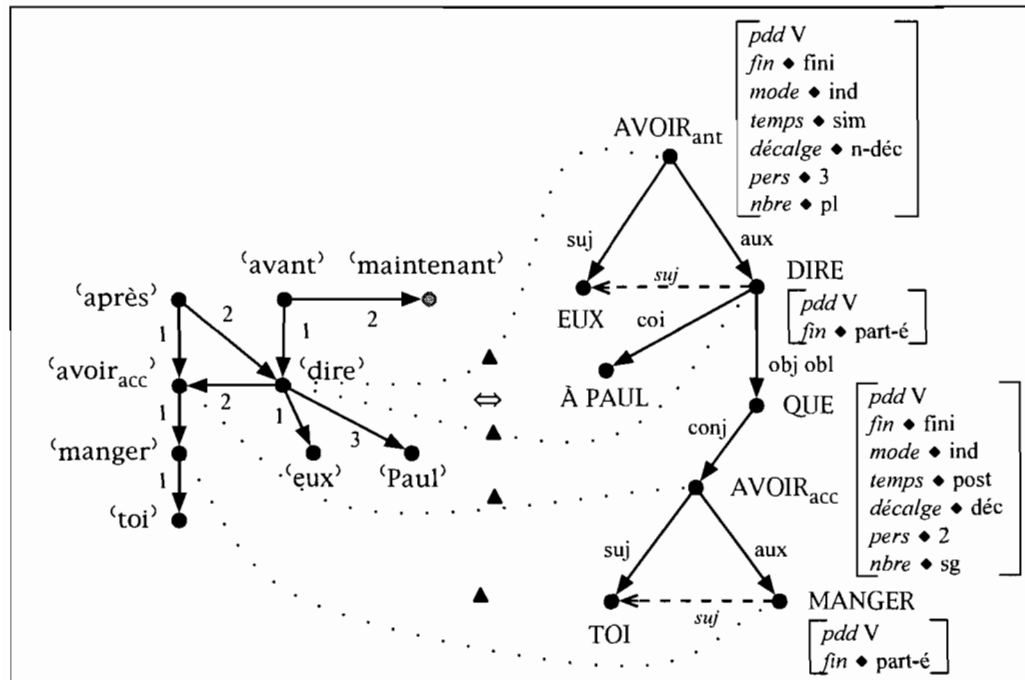


Figure 131 — La représentation complète de (175) dans $I_{\text{sém-synt}}$

3.5 Les grammies à valeur modale

Nous avons vu au début de ce chapitre (§ 1, p. 283) six sens modaux qui se décrivent dans la structure sémantique :

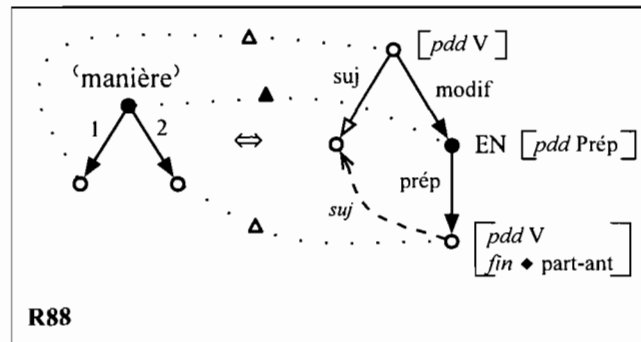
- 'X de manière Y', décrit par la règle de $G_{\text{sém}}$ R11 (p. 296).
- 'X doit Y-er', décrit par la règle R12 (p. 296).
- 'X ordonne à Y de Z-er', en R13 (p. 297).
- 'X suppose Y', en R14 (p. 297).
- 'il paraît que X', décrit par R15 (p. 298).
- 'en rétrospective, X', en R16 (p. 298).

Tous ces sens peuvent s'exprimer par des moyens lexicaux [*Le général Wanner a ordonné à ses troupes de donner l'assaut; Il paraît que le chef a pris congé;*

37. Ici encore, nous omettons certains éléments afin d'alléger la figure (cf. note 36, p. 339).

etc.]. Ces cas ne nous intéressent pas; nous allons uniquement présenter les signes qui permettent d'exprimer ces sens au moyen de grammèmes.

Le sémantème 'manière' est le signifié du grammème profond **gérondif**. Ce signe a pour signifiant la combinaison de la préposition EN et du grammème superficiel **participe-ant**:



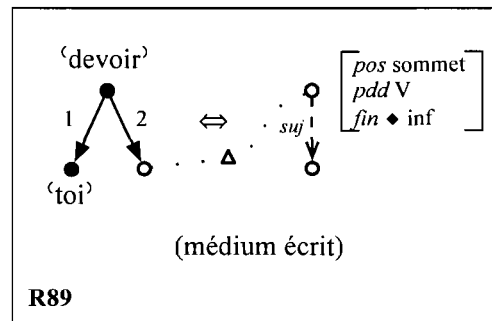
$I_{\text{sém-synt}} - p_{\text{sém-synt}}$

Figure 132 — Le gérondif dans $I_{\text{sém-synt}}$ (R88)

La diathèse de ce signe appelle quelques commentaires. Le verbe qui exprime le premier actant de 'manière' [*Il sort en claquant la porte*] se trouve modifié par la construction «*en V-ant*». Il partage obligatoirement le sujet du verbe qui réalise le second actant; ce dernier perd sa capacité à avoir un sujet (ce que représente la quasi-dépendance «suj», en pointillés). La structure qu'on obtient est semblable à celle qu'on avait pour les locutions, le passif ou le résultatif. On a vu que la copule pouvait effacer la relation «modif» (cf. § 3.2.4, p. 325). Il est donc possible de combiner la règle R88 du gérondif avec la règle R70 (p. 327) de la copule, ce qui donne une structure correspondant à des phrases du type **Il sort est en claquant la porte*. Par contre, cette structure est mal formée et sera rejetée par G_{synt} .

L'infinitif, on l'a vu plus tôt, peut exprimer une consigne uniquement quand il est le sommet de la phrase (cf. Ch. VI, § 1.1, p. 118). Pour vérifier sa position dans la structure syntaxique, nous faisons appel au trait *pos* fourni par la règle R17 (p. 301) de G_{arbre} , qui construit le sommet syntaxique. La règle R12 (p. 296) de $G_{\text{sém}}$ décrit le sémantème 'devoir' en général, et non spécifiquement en tant que signifié du signe qui nous occupe ici. Ses actants sont typés mais leur valeur reste

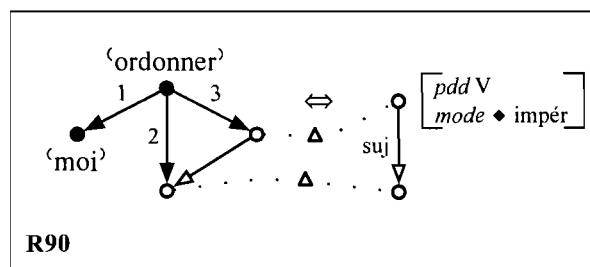
sous-spécifiée. L'infinitif de consigne, cependant, ne peut exprimer qu'une instruction s'adressant au destinataire du message. le premier actant de 'devoir' ('toi') est donc instancié dans la règle ci-dessous; il fait partie du signifié de ce signe. L'indication «médium écrit» dans cette règle fait partie de la structure rhétorique. Nous ne savons pas comment mieux encoder cette structure en GUST.



$I_{\text{sém-synt}} - P_{\text{sém-synt}}$

Figure 133 — L'infinitif de consigne $I_{\text{sém-synt}}$ (R89)

L'impératif, par contre, permet de donner un ordre non seulement à un destinataire unique [*Vas-y!*], mais aussi à un groupe de destinataires, qui peut d'ailleurs inclure le locuteur [*Allons-y!*]. Le second actant de 'ordonner' n'est donc pas instancié dans la règle ci-dessous. Par contre, son premier actant, soit la personne qui donne l'ordre, ne peut être que le locuteur dans le cas de l'impératif. Enfin, il faut absolument que le second actant sémantique corresponde au sujet du verbe. Cette règle ne construit pas la relation «sujet» (qui est polarisée en blanc); ce sont les règles de lexicalisation des verbes qui le font.



$I_{\text{sém-synt}} - P_{\text{sém-synt}}$

Figure 134 — L'impératif dans $I_{\text{sém-synt}}$ (R90)

On a vu au Chapitre VI (§ 6.4.4, p. 213) que le futur simple peut signifier une supposition de la part du locuteur [*Ils tardent. Ils seront sans doute perdus*]. Le premier actant de ‘supposer’ (‘moi’) doit donc être instancié. Ce sens est exprimé par la combinaison **postérieur** \oplus **non-décalé**. Il s’agit d’un phrasème grammématique que nous représentons comme suit :

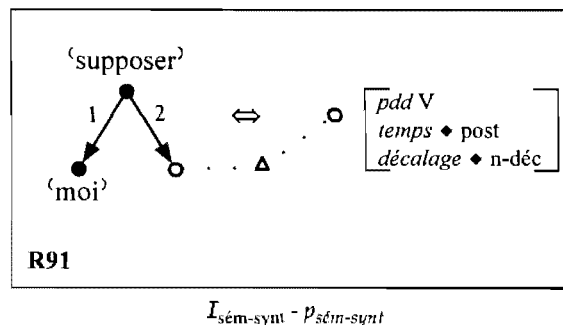


Figure 135 — Le futur de supposition dans $I_{sém-synt}$ (R91)

Nous avons aussi vu dans la même section que le conditionnel peut servir à exprimer la réserve du locuteur [*Le suspect serait entré par la fenêtre, selon les policiers*]. Nous avons décrit le sens de ce signe en $G_{sém}$ par le sémantème ‘il paraît’, dans la règle R15 (p. 298). Son signifiant est la combinaison des grammèmes **postérieur** \oplus **décalé**.

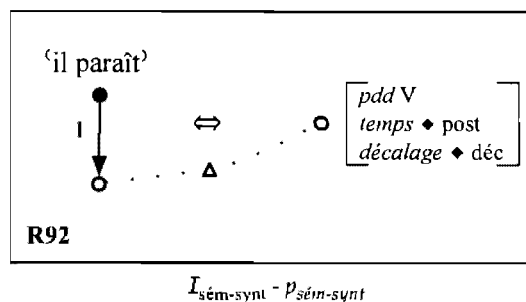


Figure 136 — Le conditionnel de réserve dans $I_{sém-synt}$ (R92)

Enfin, nous avons vu au Chapitre VI (§ 6.4.4, p. 213) que le grammème **postérieur** se combine avec AVOIR_{acc} et **non-décalé** pour exprimer un bilan rétrospectif [*L’année dernière aura été particulièrement riche en rebondissements*]. Nous avons décrit ce signe comme une collocation grammaticale. Les collocations grammaticales, à notre connaissance, n’ont jamais été discutées dans la

théorie Sens-Texte (nous n'avons pas connaissance de cette notion dans d'autres cadres théoriques non plus). Nous allons donc commencer par montrer comment se formalisent les collocations lexicales; nous pourrons ensuite mieux apprécier les parallèles avec cette collocation grammaticale.

Dans le cadre de la TST, les collocations sont décrites dans le dictionnaire par des fonctions lexicales³⁸. En GUST, comme nous l'avons dit au Chapitre VII (§ 1, p. 224), il n'y a pas de dictionnaire: tous les signes sont décrits par des règles. Kahane & Polguère (2001) ont démontré que les fonctions lexicales peuvent être définies formellement par des patrons de correspondance entre des fragments de structures sémantiques et syntaxiques. Il est donc possible de les représenter par des règles de correspondance. Prenons par exemple la collocation *peur bleue*. Son sens est 'peur intense'. Dans cette expression le lexème PEUR a été choisi librement; il y exprime son sens habituel. Il s'agit de la **base** de la collocation. Par contre, BLEU ne signifie 'intense' que dans le contexte de PEUR, et ce sens s'exprime de préférence par ce lexème dans ce contexte. Le lexème BLEU n'a donc pas été choisi tout à fait librement. C'est le **collocatif** de la collocation. Formellement, cela se représente par la règle RT6 ci-dessous, qu'on pourra comparer avec RT7:

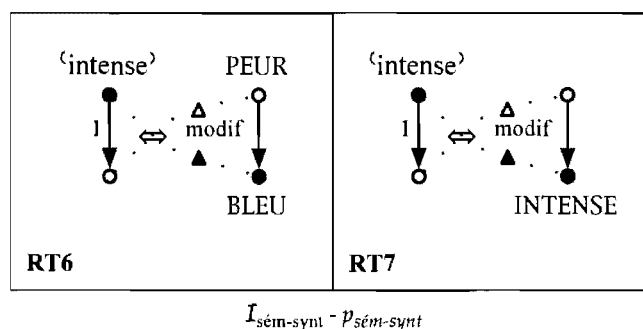


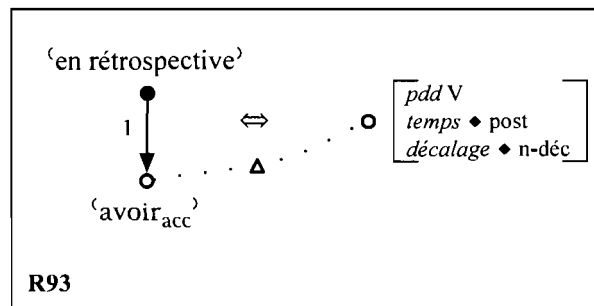
Figure 137 — Les collocations lexicales dans $I_{sém-synt}$ (RT6 et RT7)

La règle RT7 décrit la lexicalisation standard de 'intense' par l'adjectif INTENSE, comme dans l'expression *une pluie intense*, qui n'est pas une collocation³⁹. Cette règle ne tient pas compte du contexte lexical. Par contre, la règle RT6,

38. La littérature sur le sujet abonde; on pourra consulter notamment Mel'čuk (1996b).

qui décrit la collocation *peur bleue*, modélise le fait que ‘intense’ peut se lexicaliser par BLEU uniquement si cet adjectif modifie PEUR. Le sémantème qu’exprime PEUR n’est pas pertinent; ce qui compte, c’est le lexème, c’est pourquoi le sémantème n’est pas étiqueté.

Le cas du futur antérieur rétrospectif est très similaire. Son sens inclut celui de AVOIR_{acc}: ‘l’année aura été dure’ = ‘en rétrospective, l’année a été dure’. La phase accomplie est la base de la collocation. Le sens ‘en rétrospective’ ne s’exprime habituellement pas par la combinaison des grammèmes **postérieur** et **non-décalé**. Ce n’est que dans le contexte de la phase accomplie que cette combinaison de grammèmes exprime ce sens. Ils forment ensemble le collocatif. Formellement, ce signe se décrit de la façon suivante:



$I_{\text{sém-synt}} - P_{\text{sém-synt}}$

Figure 138 — Le futur antérieur rétrospectif dans $I_{\text{sém-synt}}$ (R93)

Contrairement à la règle RT6, où le sémantème correspondant à la base n’était pas pertinent, ici c’est le sémantème qui compte. On a vu plus haut (§ 3.3.3, p. 331) que ce sens peut s’exprimer soit par l’auxiliaire AVOIR_{acc}, soit par ÊTRE_{acc}. Cela n’a pas d’incidence sur la collocation:

(176) *L’année dernière aura été décevante.*

(177) *L’année dernière se sera terminée comme elle avait commencé.*

Le fait que la collocation soit contrôlée en R93 par un sémantème, et non un élément de la structure syntaxique comme en RT6, n’est pas gênant. On retrouve ce phénomène dans le domaine lexical également. Milićević (1997) montre que les

39. Justement, on lui préfère *forte pluie* ou *pluie diluvienne*, qui sont des collocations.

étiquettes sémantiques peuvent contrôler des collocations. Elle donne notamment l'exemple $Func_0(\text{'événement'}) = \text{SURVENIR, 'AVOIR LIEU', 'SE PRODUIRE'}^{40}$. Toutes les lexies qui dénotent des événements héritent de cette cooccurrence lexicale. Par exemple: *L'explosion s'est produite <a eu lieu, est survenue> à 3 h.*

4 Synthèse

Nous avons vu les règles de $G_{\text{sém}}$, G_{synt} et $I_{\text{sém-synt}}$ qui modélisent les grammèmes verbaux du français et les phénomènes connexes. En $G_{\text{sém}}$, nous avons regroupé les sens exprimés par ces grammèmes en trois types: les sens temporels, les sens phasiques ou aspectuels, et les sens modaux. En G_{synt} , nous avons décrit la combinatoire morphologique des grammèmes verbaux. Il s'agissait en fait d'une projection de la grammaire d'interface syntaxe-morphotopologie, où devrait normalement être décrite la combinatoire des morphèmes. C'est dans $I_{\text{sém-synt}}$ que nous avons décrit les grammies, c'est-à-dire les signes profonds à proprement parler, pour tous les grammèmes et quasi-grammèmes de notre modèle. Chaque règle de $I_{\text{sém-synt}}$ décrit une acception particulière.

Nous ne visions pas l'exhaustivité dans notre inventaire de ces acceptions. Notre but était surtout de décrire les principaux sens des grammèmes afin d'illustrer le traitement de la polysémie grammaticale dans GUST. En particulier, nous avons dû laisser de côté les signes qui ne se laissaient pas représenter par une mise en correspondance d'éléments de la structure sémantique et de la structure syntaxique. Les signifiés des signes grammaticaux verbaux du français peuvent faire intervenir les structures communicative, rhétorique et référentielle, que nous ne savons pas décrire en GUST. Leurs signifiants peuvent aussi relever d'autres structures que la structure syntaxique; c'est le cas notamment du présent d'injonction [*On se calme!*], qui s'exprime par un prosodème.

40. La fonction lexicale $Func_0$ associe à un mot-clé le verbe support qui prend ce mot-clé comme sujet. Par exemple, $Func_0(\text{PLUIE}) = \text{TOMBER [la pluie tombe]}$.

Même en laissant de côté ces signes, l'étude de la flexion verbale reste particulièrement intéressante puisqu'elle met en jeu plusieurs types de signes. Les grammèmes syntaxiques n'expriment pas de sens et ne sont donc mis en correspondance avec aucun objet sémantique. Les grammèmes de décalage, dans leur acception de base, sont mis en relation avec des objets sémantiques qui ne sont pas saturés dans les règles. Enfin, les grammèmes verbaux sont aussi impliqués dans des phrasèmes grammémiques et des collocations grammaticales.

On a pu également observer la dynamique de l'articulation des modules de GUST. On a vu par exemple que G_{synt} permet de choisir le bon site d'ancrage pour les grammèmes ou les arcs syntaxiques quand un sémantème a plus d'un nœud correspondant en syntaxe. Nous avons illustré également comment G_{synt} peut forcer l'application de règles de $I_{\text{sém-synt}}$, même après que ce module ait complété la saturation en $p_{\text{sém-synt}}$ de tous les objets. On a pu voir aussi que G_{synt} peut rejeter des structures mal formées construites par $I_{\text{sém-synt}}$. Ce mécanisme peut donner l'impression d'une architecture de type «génération-filtrage», où un premier module génère des structures qu'un second module filtre ensuite. Ce type de modèle pose des problèmes d'implémentation et sont, d'un point de vue cognitif, douteux. Mais l'architecture de GUST n'est pas de ce type. Les modules $I_{\text{sém-synt}}$ et G_{synt} ne sont pas appliqués de façon séquentielle. Les règles de G_{synt} s'appliquent en même temps que celles de $I_{\text{sém-synt}}$. Ensemble, tous les modules de GUST ne forment en fait qu'une seule grammaire, grâce au mécanisme d'articulation décrit au Chapitre VII (§ 5, p. 236).

Cette architecture permet de décrire les phénomènes linguistiques par petits fragments. L'écriture des règles en est beaucoup simplifiée, puisque chaque règle ne doit s'occuper que d'un élément atomique de la grammaire. Il en résulte des règles d'une extrême simplicité, comme on a pu le constater, et c'est là un des points forts de GUST.

IX. Conclusion

Le but de cette thèse était double. D'une part, nous voulions décrire la flexion verbale en français, et plus particulièrement celle en temps grammatical dans l'interface sémantique-syntaxe, en mettant l'accent sur l'étude des signes grammaticaux et de leur organisation. D'autre part, nous cherchions, à travers ce phénomène linguistique central dans la grammaire du français, à vérifier l'utilité de GUST pour la construction de modèles formels de la langue.

Pour atteindre notre premier objectif, nous avons d'abord présenté les notions fondamentales nécessaires à l'étude des signes flexionnels. Au cœur de ces notions se trouve le grammème, conçu comme une entité du même niveau d'abstraction que le vocable. Il s'agit d'un élément des représentations syntaxiques profondes (pour les grammèmes profonds) ou de surface (pour les grammèmes superficiels). Nous nous sommes plus particulièrement intéressé aux grammèmes superficiels, puisque GUST ne comprend pas de niveau syntaxique profond.

Les grammèmes peuvent être polysémiques; à chaque acception correspond une grammie, c'est-à-dire un signe profond de nature grammaticale. Nous avons proposé des critères sémantiques et syntaxiques, inspirés de la lexicographie, pour identifier parmi les diverses grammies d'un même grammème celle qui constitue la grammie de base. Notre méthodologie pour l'étude des systèmes flexionnels s'appuie fortement sur ces critères, puisque ce sont les grammies de base qui guident le regroupement des grammèmes en catégories flexionnelles présentant une certaine homogénéité à la fois dans la combinatoire et dans le sens des grammèmes qu'elles contiennent.

Toutefois, la seule étude de la flexion verbale en français n'a pas permis de vérifier pleinement ni d'infirmer la validité de notre méthodologie. Nous ne savons pas dans quelle mesure les critères proposés pour l'identification de la grammie de base d'un grammème s'appliquent à l'étude d'autres catégories flexionnelles en français ou dans d'autres langues. Par ailleurs, bien qu'en français les grammèmes

temporels se laissent élégamment regrouper en catégories flexionnelles quand on ne considère que leur grammaire de base, il n'est pas certain que cette stratégie puisse fonctionner pour toutes les catégories flexionnelles qu'on retrouve dans les langues naturelles. Notre méthodologie, développée de façon plus ou moins intuitive, mériterait d'être testée sur un plus large spectre de phénomènes linguistiques.

Néanmoins, cette méthodologie nous a bien servi pour la description de la flexion verbale en français et nous a permis d'aboutir à un système cohérent où l'accent est mis sur les signes dans leur intégralité. Une des principales critiques que nous avons faites des descriptions existantes est qu'elles insistent souvent trop sur une seule des composantes du signe linguistique. Forcément, on aboutit ainsi à des modèles du sens, de la forme ou de la combinatoire des signes, mais pas à une description du sous-système sémiotique dans son entier.

Au niveau descriptif, nous avons utilisé les formes surcomposées pour prouver l'hypothèse de Benveniste (1959) selon laquelle l'auxiliaire **AVOIR** peut exprimer soit l'antériorité, soit l'accompli, ajoutant du poids aux arguments déjà proposés par lui-même et d'autres linguistes. Nous avons poussé l'hypothèse encore plus loin en suggérant que **AVOIR** n'est pas vague, mais bien polysémique, c'est-à-dire qu'il existe en français deux auxiliaires distincts, **AVOIR_{ant}** et **AVOIR_{acc}**. Le premier sert de signifiant à un grammème profond de temps, alors que le second fait partie des marqueurs de phase aspectuelle.

Les marqueurs de phase aspectuelle dénotent les périodes qui se trouvent (immédiatement) avant, pendant ou (immédiatement) après un fait. Il s'agit des lexies **AVOIR_{acc}**, **«EN TRAIN»**, **ALLER**, **«SUR LE POINT»**, et **VENIR**. Ces lexies ne constituent pas une catégorie flexionnelle, notamment parce qu'elles ne sont pas mutuellement exclusives. Elles forment de nouvelles bases qui peuvent elles-mêmes être fléchies (sauf pour les locutions prépositionnelles **«EN TRAIN»** et **«SUR LE POINT»**, puisque les adverbes ne sont pas fléchis en français). En particulier, elles se combinent aux grammèmes temporels.

C'est la description des grammèmes temporels qui était au cœur de la partie descriptive de cette thèse. L'hypothèse que nous défendons est que les relations temporelles sont représentées en français au moyen de deux catégories flexionnelles qui se combinent :

		Décalage	
		non-décalé	décalé
Temps	antérieur	<i>a aimé / aima</i>	<i>avait aimé</i>
	simultané	<i>aime</i>	<i>aimait</i>
	postérieur	<i>aimera</i>	<i>aimerait</i>

Tableau XXVIII — Deux catégories flexionnelles complémentaires

- 1) La catégorie de décalage regroupe deux grammèmes situant un point de repère par rapport au moment d'énonciation :
 - Le grammème **non-décalé** signifie que le point de repère est 'maintenant' ou un fait dans le futur. Son signifiant est nul.
 - Le grammème **décalé** indique que le repère est un fait dans le passé. Il s'exprime par le suffixe **-AI-**.
- 2) La catégorie de temps regroupe trois grammèmes profonds situant les faits par rapport à ce point de repère :
 - Le grammème **simultané** signifie que le fait dénoté par le verbe est simultané à son point de repère. Il s'exprime par un suffixe zéro.
 - Le grammème profond **antérieur** indique que le fait est antérieur à son point de repère. Son signifiant est l'auxiliaire **AVOIR_{ant}** (ou **ÊTRE_{ant}**). Dans un registre littéraire, il s'exprime conjointement avec le grammème **non-décalé** par le suffixe du passé simple.
 - Le grammème **postérieur** pose le fait comme postérieur à son point de repère. Il s'exprime par le suffixe **-R-**.

L'idée sous-jacente à notre modèle n'est pas neuve; on a depuis longtemps remarqué la dichotomie «temps absolus» ~ «temps relatifs». Cependant, ce fait linguistique en français, à notre connaissance, n'avait jamais été formulé en termes

de catégories flexionnelles *complémentaires* dont les grammèmes se combinent, mais toujours en termes de catégories *concurrentes* (ou de partitions sémantiques d'une seule catégorie flexionnelle). La conception traditionnelle ne met pas assez en relief le caractère compositionnel des signes grammaticaux. Dans notre modèle, aucune forme verbale n'exprime qu'un seul grammème temporel. Toutes expriment un grammème de temps *et* un grammème de décalage. Nous aimerions proposer une terminologie qui reflète cela: «présent (non-décalé)» (pour le «présent»), «passé (non-décalé)» (plutôt que «passé composé»), «futur (non-décalé)» (plutôt que «futur simple»), «présent décalé» (plutôt que «imparfait»), «passé décalé» (plutôt que «plus-que-parfait») et «futur décalé» (plutôt que «conditionnel»). Dans le cas du passé simple, nous parlerions de «passé (non-décalé) synthétique».

Nous n'avons pas cherché à vérifier de façon systématique à quel point notre modèle est utile pour la description de la flexion verbale dans d'autres langues que le français. Nous avons toutefois tenté une expérience préliminaire dans le cadre du projet *Marquis* à l'Université Pompeu Fabra, dont l'objectif était de construire un système informatique qui génère automatiquement des bulletins sur la qualité de l'air dans huit langues européennes typologiquement diverses (cf. Wanner et al. 2007). Étant donné la nature multilingue de ce projet, nous avons adopté une architecture particulière pour les grammaires, où la plupart des règles étaient partagées par plusieurs langues (cf. Lareau & Wanner 2007). Pour les langues romanes que nous avons à décrire (le catalan, l'espagnol, le français et le portugais), nous avons réutilisé le même modèle pour la flexion verbale. Nous avons alors pu constater que ces quatre langues montrent la même organisation à ce niveau. La distinction stricte que nous faisons entre AVOIR_{ant} et AVOIR_{acc} en français trouve un écho direct dans les trois autres langues, où il existe deux formes distinctes pour exprimer les sens correspondants. Nous avons toutefois noté que la sélection du point de repère ne se fait pas forcément de manière identique dans toutes ces langues. Par exemple, l'imparfait dans *Ce matin, la concentration d'ozone était de 25 µg/m³* se traduit plutôt par le passé *va ser* en catalan, et non par l'imparfait **era*. Ce fait sug-

gère que la saillance communicative des circonstanciels n'est pas la même d'une langue à l'autre, et que cela influence le choix du point de repère pour la flexion verbale. Il y a là matière à réflexion.

Notre second objectif, dans cette thèse, était de vérifier l'utilité de GUST pour la description des langues, tout en développant le formalisme pour le rendre plus approprié à cette tâche. Nos principaux apports à ce formalisme ont été la représentation des structures en termes d'objets et de fonctions sur ces objets, ainsi que la modélisation des décompositions sémantiques.

La formalisation de notre modèle descriptif de la flexion verbale nous a donné l'occasion d'observer le mécanisme d'articulation des modules de GUST. On a vu comment la grammaire de bonne formation syntaxique fonctionne de paire avec le module d'interface sémantique-syntaxe pour générer des structures bien formées. On a pu voir aussi que ce formalisme permet d'exprimer d'une façon simple et assez intuitive les règles qui modélisent des phénomènes relativement complexes, notamment grâce à une architecture hautement modulaire où les problèmes sont éclatés au maximum. L'exercice nous a également permis de vérifier l'utilité du mécanisme de décomposition des sens linguistiques pour le typage des actants des prédicats sémantiques.

Le choix de la flexion verbale comme phénomène linguistique pour vérifier l'utilité de GUST s'est avéré judicieux, puisqu'elle fait intervenir plusieurs types de signes. Nous avons pu décrire de façon élégante l'accord verbal dans la phrase simple et d'autres grammèmes syntaxiques, la montée du sujet et d'autres opérations syntaxiques, les grammèmes de décalage qui, dans leur acception de base, ne saturent aucun objet de la représentation sémantique, ou encore les phrasèmes grammémiques et les collocations grammaticales. En ce sens, l'expérience démontre que GUST est en principe approprié pour décrire une variété de phénomènes linguistiques, du moins dans l'interface sémantique-syntaxe.

Par contre, nous avons dû laisser de côté certains faits linguistiques que nous ne savions pas décrire en GUST. Nous nous sommes limité dans cette thèse aux

signes mettant en jeu les structures sémantique et syntaxique. Or, nous avons observé qu'il existe des signes grammaticaux verbaux dont le signifié relève de la structure communicative, rhétorique ou référentielle, ou dont le signifiant relève de la structure prosodique. Nous ne savons pas comment formaliser ces structures en GUST. Cependant, il ne s'agit pas d'une limitation intrinsèque à cette théorie. En fait, Kahane (2004, 2005) et Kahane & Lareau (2005a) ont démontré que GUST (ou le formalisme plus général des grammaires d'unification polarisées) pouvait modéliser des structures substantiellement différentes. Il s'agit maintenant de chercher la meilleure façon d'encoder ces structures afin de les intégrer au modèle.

Enfin, puisque l'écrasante majorité des phrases en français comportent au moins un verbe, la flexion verbale est un élément clé de la grammaire de cette langue. Il nous paraît donc sensé de commencer la description du français en GUST par ce phénomène. Nous souhaitons poursuivre le travail commencé et étendre la couverture de la grammaire à d'autres phénomènes. Nous espérons toutefois que cette thèse aura fourni de nouvelles idées utiles non seulement à la description du français mais également à celle d'autres langues, et que le formalisme développé se révélera un outil puissant pour la modélisation des langues naturelles.

Index des auteurs

Abeillé	25, 134, 314
Beck	67
Bello	53
Benveniste	77, 159, 161, 350
Bescherelle	120, 178, 193
Beysade	48
Binnick	70
Blanche-Benveniste	25
Boas	51
Bohnet	224
Bres	199
Bresnan	224
Candito	248, 311
Carruthers	33, 34, 145, 146, 147, 148, 156, 157, 158, 163
Chu	175
Clas	46, 53, 55, 285
Comrie	73, 80, 108, 181, 182, 197, 202, 203
Cornu	25, 150, 153, 154, 156, 157, 159
Creissels	125, 126, 127, 128
Curat	159
Damourette	151, 206
de Saussure	255
Destutt de Tracy	78, 79
Dobrovie-Sorin	48
Dowty	60
Duchier	230
El-Kassas	208, 254
Foulet	155, 157
Garey	181
Gerdes	230, 254
Girard	78, 79
Godard	25, 134, 314
Gosselin	25, 60, 61, 106, 143, 170, 181, 197, 286, 288
Grevisse	73, 74, 118, 121, 122, 123, 137, 142, 152, 153, 155, 163, 193, 198
Guillaume	139
Haillet	201
Hill	145, 147, 148
Imbs	73, 80, 159, 162, 179, 197, 210
Iordanskaja	42, 245, 311, 312, 313, 314, 315, 316
Jakobson	51, 54
Jolivet	25, 145, 150, 152, 157
Joshi	224
Kahane	35, 36, 46, 49, 117, 132, 133, 223, 224, 225, 226, 227, 230, 231, 237, 238, ... 242, 243, 247, 249, 254, 255, 269, 274, 317, 323, 324, 325, 326, 331, 345, 354

Kahn	206
Kamp	185, 186, 187, 191
Kayne	134
Labeau	193, 199
Lareau	223, 224, 230, 237, 254, 269, 352, 354
Larrivée	199
Laurendeau	25
Le Goffic	25, 26, 142, 193, 199
Leeman-Bouix	135, 139, 142, 159, 166, 170
Lefevre	245
Martinet	46, 63, 64, 67, 84, 91, 137, 184, 196, 207
Mazon	182
Mel'čuk	25, 35, 39, 40, 41, 42, 44, 46, 47, 48, 50, 51, 53, 54, 55, 65, 66, 67, 117, 120, 124, 125, 132, 133, 136, 137, 166, 202, 222, 224, 242, 245, 255, 267, 268, 269, 274, 277, 285, 288, 311, 312, 313, 314, 315, 316, 322, 323, 345
Menanteau	89, 90, 139, 140
Milićević	35, 241, 269, 285, 346
Molendijk	179, 180, 185, 187, 191
Morin	25
Mourelatos	285, 291, 295
Nasr	230
Partee	60, 80
Pellat	29, 33, 51, 52, 56, 73, 75, 121, 138, 142, 169, 170, 193
Perrier	230
Pertsov	245
Pichon	151, 206
Pinchon	135, 165, 166, 170, 173
Polguère	35, 41, 44, 46, 47, 53, 55, 226, 237, 241, 268, 269, 270, 277, 285, 323, 345
Pollard	224
Psyché	241
Reichenbach	60, 187, 188, 189, 190, 191, 192, 195, 203, 204, 208
Riegel	29, 33, 51, 52, 56, 73, 75, 121, 138, 142, 169, 170, 193
Rioul	29, 33, 51, 52, 56, 73, 75, 121, 138, 142, 169, 170, 193
Rohrer	185, 186, 187, 191
Sag	224
Sapir	51
Te Winkel	192, 193
Tesnière	117, 157, 162
Thater	230
Touratier	25, 63, 64, 90, 91, 95, 135, 136, 196, 207
Van Den Eynde	25
Vendler	175, 181, 285, 286
Verkuyl	192
Vet	25, 71, 184, 206, 210
Vetters	71, 77, 79, 181, 182, 183, 187, 189, 191, 209, 286

Wagner	135, 165, 166, 170, 173
Wanner	224, 269, 352
Weinrich	210
Wierzbicka	269
Wilmet	96
Yoo	254
Žolkovskij	35

Index des notions

A

accompli	76, 159, 164, 166, 220, 281, 289, 303
accomplissement	181
accord	220, 306, 310, 312, 314, 319
achèvement	181
actif	133
activité	181
Aktionsart	181
aller	28, 164, 170, 173, 220, 281, 289, 303, 311, 331
analyse	263
antérieur	159, 178, 197, 200, 205, 219, 220, 221, 281, 286
arc sémantique (type d'objet)	239
arc syntaxique (type d'objet)	245
articulation (des grammaires)	249, 250, 251, 260, 263, 264
articulation $G_{\text{sém}}-I_{\text{sém-synt}}$	260
articulation $I_{\text{sém-synt}}-G_{\text{synt}}$	260
aspect	164, 181
aspect grammatical	181
auxiliaire	28, 145, 301, 311
avoir . 28, 77, 145, 158, 164, 166, 178, 197, 200, 220, 281, 289, 303, 305, 311, 331	

B

base (d'une collocation)	345
bonne formation (des représentations)	240, 248

C

catégorie d'accord	220
catégorie flexionnelle	46, 50, 62
cible (fonction)	239, 245
collocatif	345
collocation	345
collocation grammaticale	219, 344
combinaison (de règles)	233
combinaison partielle	236
combinaison totale	235
combinatoire	63, 64
concordance des temps	212
conditionnel	142, 200, 202, 217, 218, 222, 283, 309, 344
conditionnel passé	176
conditionnel surcomposé	150
contrainte de déclarativité	233

contrainte de dynamisme	232
contrainte de finalité	233
contrainte de monotonie	232
copule	313
correspondance	255, 317
critère de l'exclusion mutuelle	64
critère de l'inclusion de sens	55
critère de l'interprétation spontanée	54
critère de la polyvalence lexicale	59
critère de la polyvalence syntaxique	58
critère de la similitude de combinatoire	64
critère du sommet syntaxique	58

D

décalage	206, 207, 217, 219, 221, 281, 289, 310, 334, 335
décalé	217, 219, 221, 281, 289, 309, 336, 344
déclarativité (contrainte)	233
décomposition sémantique	60, 267
dérivation	51
dynamisme (contrainte)	232

E

en train	28, 164, 169, 220, 281, 289, 303
état	181
étiquette (fonction)	239, 243, 244, 245
être	28, 134, 303, 305, 311, 331
évidentialité	222

F

finalité (contrainte)	233
fini	130, 131, 221, 307
finitude	116, 117, 221, 283, 301, 307, 342
flexion	51
fonction	227
fonction d'étiquetage	228
fonction de polarisation	228
fonction structurante	228, 243
forme surcomposée	144, 145, 176, 220
futur	352
futur antérieur	176
futur antérieur surcomposé	152
futur décalé	352
futur inéluctable	214
futur périphrastique	176

futur simple 170, 198, 202, 205, 217, 218, 222, 283, 344

G

générique (fonction) 270, 284
 genre 220, 310
 gérondif 129, 132, 307, 342
 gérondif passé 176
 grammaire d'arbres 260, 299
 grammaire d'interface 255
 grammaire d'interface sémantique-syntaxe 256, 318
 grammaire d'interface syntaxe-morphotopologie 298, 301
 grammaire d'unification Sens-Texte 223, 224
 grammaire de bonne formation 237
 grammaire de bonne formation sémantique 239, 270, 283, 284, 298
 grammaire de bonne formation syntaxique 242
 grammème 46, 49, 53, 264, 307
 grammème profond 50, 307
 grammème superficiel 50, 307
 grammie 49, 317
 grammie de base 53
 graphe de correspondance 257

H

hôte (fonction) 244

I

imparfait 178, 199, 200, 202, 212, 214, 215, 216, 283
 impératif 141, 144, 221, 304, 305, 308
 impératif passé 176
 imperfectif 181, 182
 indicatif 135, 144, 221, 304, 305, 306, 308, 310
 infinitif 118, 120, 131, 221, 307
 infinitif passé 176
 infinitif passé surcomposé 154

L

langue 35
 locution prépositionnelle 328

M

masdar 117
 mode 75, 134, 221, 283, 304, 305, 306, 308, 310, 343
 modèle Sens-Texte 36

module	224, 225
module de correspondance	43
monotonie (contrainte)	232

N

nœud sémantique (type d'objet)	239, 270
nœud syntaxique (type d'objet)	243
nombre	220, 309, 310, 319, 320
non-décalé	217, 219, 221, 281, 289, 335, 344

O

objet	227
objet grammématique	243
obj-gramm (type d'objet)	244, 260

P

p _{arbre} (fonction)	299
participe	121, 122, 307
participe passé	121, 124, 310, 320, 322
participe présent	121, 123, 322
participe présent composé	176
participe surcomposé	154
participe-ant	123, 131, 221, 307
participe-é	123, 124, 132, 221, 307
passé	352
passé antérieur	155, 162
passé antérieur surcomposé	152
passé composé	155, 159, 166, 176, 178, 197, 202, 205, 212, 219
passé décalé	352
passé récent	176, 214, 281
passé simple	155, 178, 197, 202, 205, 212, 219
passé surcomposé	146, 155
passé synthétique	352
passif	133, 134, 161, 313
p _{dd} (fonction)	243
perfectif	181, 182
personne	220, 306, 309, 312, 314, 320
phase aspectuelle	164, 165, 173, 220, 281, 289, 303, 327
phrasème grammématique	67, 344
pluriel	320
plus-que-parfait	164, 176, 200, 202
plus-que-parfait surcomposé	150, 155
point d'articulation	251, 260
polarisation	239, 245

polarité	230, 249
polarité d'articulation	252
polarité de construction	252
polarité neutre	232
polarité principale/secondaire (d'une figure).	249, 282
postérieur	205, 217, 219, 221, 281, 286, 309, 334, 344
présent	196, 202, 214, 215, 352
présent décalé	352
primitif sémantique	272
produit (des polarités)	231
progressif	169, 281
prospectif	170, 281
P _{sém} (fonction)	239, 284
P _{sém-synt} (fonction)	318
P _{synt} (fonction)	245, 299

Q

quasi-dépendance	324
----------------------------	-----

R

réduction (fonction)	270, 284
réfléchi	133, 161
règle de correspondance	43
règle grammémique	247
règle lexicale	246
règle sagittale	246
relation syntaxique	311
représentation linguistique	36, 38
représentation morphologique	42
représentation phonologique	42
représentation sémantique	38, 239, 240, 242
représentation syntaxique de surface	42
représentation syntaxique profonde	40
résultatif	124, 129, 293, 313, 328

S

sens grammatical	60
sens modal	295
signe asémantique	319
signe linguistique	255
signe profond	317
simultané	205, 217, 219, 221, 281, 286, 334
singulier	320
source (fonction)	239, 245

structure anaphorique profonde	41
structure communicative	40, 242, 283
structure élémentaire	226
structure initiale	227, 245, 300
structure polarisée	229
structure référentielle	40, 283
structure rhétorique	40, 283
structure sémantico-communicative	40
structure sémantique	38, 283
structure syntaxico-communicative profonde	41
structure syntaxico-prosodique profonde	41
structure syntaxique	283
structure syntaxique de surface	42
structure syntaxique profonde	41
subjonctif	137, 144, 221, 283, 304, 305, 306, 308, 310
subjonctif passé	176
subjonctif passé surcomposé	153
subjonctif plus-que-parfait surcomposé	153
sujet	316
sur le point	28, 172, 173, 220, 281, 289, 303
surcomposé	32
synthèse	263

T

TAG	247
temps	177, 195, 202, 205, 219, 221, 281, 286, 305, 309, 334
temps absolu	77, 202
temps absolu-relatif	203
temps relatif	77, 202
théorie Sens-Texte	35, 223
traits de syntactique	301
typage des actants sémantiques	241

U

unification (d'objets)	233
----------------------------------	-----

V

valence sémantique	239
venir	28, 172, 173, 220, 281, 289, 303, 311, 331
voix	122, 132, 161, 283

X

XML (format)	xxiv
------------------------	------

Bibliographie

- Abeillé, A. & Godard, D. (2002). The syntactic structure of French auxiliaries. *Language*, 78(3): 404–452.
- Arnauld, A. & Lancelot, C. (1993 [1660]). *Grammaire générale et raisonnée de Port-Royal*. Slatkine Reprints, Genève.
- Beck, D. (2007). Morphological phrasemes in Totonacan inflection. In *Proceedings of MTT 2007*, Klagenfurt.
- Bello, A. (1982 [1847]). *Gramática de la lengua castellana*. Edaf, Madrid.
- Benveniste, E. (1959). Les relations de temps dans le verbe français. *Bulletin de la Société de linguistique de Paris*, 54(1): 69–82. Réédité dans Benveniste, E. (1966). *Problèmes de linguistique générale*. Gallimard, Paris.
- Bescherelle. *La conjugaison pour tous*. (1997). Edelsa, Paris.
- Bescherelle. *L'art de conjuguer: dictionnaire de 12 000 verbes*. (1998). Hurtubise HMH, Montréal.
- N.B.:** Le *Bescherelle* n'est pas édité par la même compagnie au Canada et en France. Le contenu des deux éditions est toutefois sensiblement le même, et puisque nous renvoyons aux numéros d'articles plutôt qu'aux numéros de pages, le lecteur pourra retrouver le contenu cité dans l'édition qu'il a à sa disposition.
- Beyssade, C. and Dobrovie-Sorin, C. (2005). *Définir les indéfinis*. CNRS Éditions, Paris.
- Binnick, R. I. (2001). *A Bibliography of Tense, Verbal Aspect, Aktionsart, and Related Area: 6600 Works*. University of Toronto.
- Boas, F. (1938). Language. *General Anthropology*.
- Bohnet, B., Langjahr, A. & Wanner, L. (2000). A Development Environment for an MTT-Based Sentence Generator. In *Proceedings of the First International Conference on Natural Language Generation*, Mitzpe Ramon.
- Bohnet, B. & Wanner, L. (2001). On using a parallel graph rewriting formalism in generation. In *Proceedings of the 8th European workshop on Natural Language Generation*, Toulouse: 1–11.
- Bres, J. (2005). *L'imparfait dit narratif*. CNRS Éditions, Paris.

- Bresnan, J. (2001). *Lexical-functional syntax*. Malden / Blackwell, Oxford.
- Candito, M.-H. (1999). *Organisation modulaire et paramétrable de grammaires électroniques lexicalisées. Application au français et à l'italien*. Thèse de doctorat, Université Paris 7.
- Carruthers, J. (1993). *The formes surcomposées: The Discourse Function and Linguistic Status of a Rare Form in Contemporary Spoken French*. Thèse de doctorat, University of Cambridge.
- Carruthers, J. (1994). The *passé surcomposé régional*: towards a definition of its function in contemporary spoken French. *Journal of French Language Studies*, 4(2): 171–190.
- Carruthers, J. (1996). The *passé surcomposé général*: on the relationship between a rare tense and discourse organization. *Romance Philology*, 50(2): 183–200.
- Carruthers, J. (1998). *Surcomposé général et surcomposé régional: deux formes distinctes?* In *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza*, Niemeyer, Tübingen: 143–154.
- Carruthers, J. (1999). A problem in sociolinguistic methodology: investigating a rare syntactic form. *Journal of French Language Studies*, 9: 1–24.
- Chu, X. (2008). *Les verbes modaux du français*. Ophrys, Paris.
- Comrie, B. (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge University Press, Cambridge.
- Comrie, B. (1985). *Tense*. Cambridge University Press, Cambridge.
- Cornu, M. (1953). *Les formes surcomposées*. Francke, Bern.
- Creissels, D. (2000). L'emploi résultatif de *être + participe passé* en français. In Carlier, A. et al. (dir.). *Cahiers Chronos 6: Passé et parfait*. Rodopi, Amsterdam: 133–142.
- Curat, H. (1991). *Morphologie verbale et référence temporelle en français moderne: Essai de sémantique grammaticale*. Droz, Genève.

- Damourette, J. & Pichon, É. (1911-1950). *Des mots à la pensée. Essai de grammaire de la langue française*. D'Artrey, Paris.
- De Saussure, F. (1931). *Cours de linguistique générale*. Payot, Paris.
- Destutt de Tracy, A.-L.-C. (1803). *Éléments d'idéologie*. Courcier, Paris.
- Dictionnaire de l'Académie française*. (1878). 7^e édition. Institut de France, Paris.
- Dowty, D. (1972). *Studies in the Logic of Verb Aspect and Time Reference*. Thèse de doctorat, Austin, University of Texas.
- Duchier, D. & Thater, S. (1999). Parsing with tree descriptions: a constraint-based approach. In *Proceedings of NLULP 1999*, Las Cruces.
- El-Kassas, D. (2005). *Une étude contrastive de l'arabe et du français dans une perspective de génération multilingue*. Thèse de doctorat, Université Paris 7.
- Foulet, L. (1925). Le développement des formes surcomposées. *Romania*, 51: 203–252.
- Garey, H. B. (1957). Verbal aspect in French. *Language*, 33: 91–110.
- Gerdes, K. (2001). TAG and Topology - Problems and Proposals for German. In *Proceedings of FGMOL 2001*, Helsinki.
- Gerdes, K. (2002). *Topologie et grammaires formelles de l'allemand*. Thèse de doctorat, Paris, Université Paris 7.
- Gerdes, K. & Kahane, S. (2001). Word Order in German: A Formal Dependency Grammar Using a Topological Hierarchy. In *Proceedings of ACL 2001*, Toulouse.
- Gerdes, K. & Kahane, S. (2006). L'amas verbal au cœur d'une modélisation topologique de l'ordre des mots. *Linguisticae Investigationes*, 29(1): 75–89.
- Girard (1747). *Vrais principes de la langue française*. Paris.
- Gosselin, L. (1996). *Sémantique de la temporalité en français. Un modèle calculatoire et cognitif du temps et de l'aspect*. Duculot, Louvain-la-Neuve.
- Grevisse, M. [édition revue par A. Goosse,] (1993). *Le bon usage* (13^e éd.). Duculot, Louvain-la-Neuve.

- Haillet, P. P. (2002). *Le conditionnel en français: une approche polyphonique*. Ophrys, Paris.
- Hill, J. K. (1984). *À la recherche de temps perdus: the double compound forms of the verb in present-day French*. *Word*, 35(1): 89–112.
- Imbs, P. (1960). *L'emploi des temps verbaux en français moderne*. Klincksieck, Paris.
- Iordanskaja, L. & Mel'čuk, I. A. (à paraître). Establishing an Inventory of Surface-Syntactic Relations: Valency-Controlled Surface-Syntactic Dependents of the Verb in French. In Mel'čuk, I. A. & Polguère, A. (dir.). (2008). Une version antérieure et plus courte de cet article a été publiée dans Iomdin, L. L. & Krysin, L. P. (dir.). (2000). *Slovo v tekste i v slovare. Sbornik statej k semidesjatiletiju akademika Ju. D. Apresjana*. Jazyki russkoj kul'tury, Moscou: 391–433.
- Jakobson, R. (1959). Boas' view of grammatical meaning. In *The Anthropology of Franz Boas*. *American Anthropologist*, 61(5-2). Traduit et réédité dans Jakobson (1963).
- Jakobson, R. (1963). *Essais de linguistique générale*. Éditions de Minuit, Paris.
- Jolivet, R. (1984). L'acceptabilité des formes verbales surcomposées. *Le Français Moderne*, 52: 159–176.
- Joshi, A., Levy, L. & Takahashi, L. (1975). Tree Adjunct Grammars. *Journal of the Computer and System Sciences*, 10(1): 136–163.
- Kahane, S. (1998). Le calcul des voix grammaticales. *Bulletin de la Société de linguistique de Paris*, 93(1): 325–348.
- Kahane, S. (2000a). Extractions dans une grammaire de dépendance à bulles, *Traitement Automatique des Langues (T.A.L.)*, 41(1): 187–216.
- Kahane, S. (2000b). Des grammaires formelles pour définir une correspondance. In *Actes de TALN 2000*, Lausanne: 197–206.
- Kahane, S. (2001). Grammaires de dépendance formelles et théorie Sens-Texte. In *Actes de TALN 2001*, Tours: vol.2, 17–76.

- Kahane, S. (2002). *Grammaire d'unification sens-texte: vers un modèle mathématique articulé de la langue*. Document de synthèse pour l'habilitation à diriger des recherches, Université Paris 7.
- Kahane, S. (2003a). The Meaning-Text Theory. In Agel V. *et al.* (dir.) *Dependenz und Valenz: Ein internationales Handbuch der zeitgenössischen Forschung / Dependency and Valency: An International Handbook of Contemporary Research*. Walter de Gruyter, Berlin / New York: vol. 1, 546–570.
- Kahane, S. (2003b). Les signes grammaticaux dans l'interface sémantique-syntaxe d'une grammaire d'unification. In *Actes de TALN 2003*, Batz-sur-mer: 175–184.
- Kahane, S. (2003c). On the status of the deep syntactic structure. In *Proceedings of MTT 2003*, Paris.
- Kahane, S. (2004). Grammaires d'unification polarisées. In *Actes de TALN 2004*, Fès.
- Kahane, S. (2005), Structure des représentations logiques, polarisation et sous-spécification. In *Actes de TALN 2005*, Dourdan: 153–162.
- Kahane, S. & Lareau, F. (2005a). Grammaire d'Unification Sens-Texte: modularité et polarisation. In *Actes de TALN 2005*, Dourdan: 23–32.
- Kahane, S. & Lareau, F. (2005b). Meaning-Text Unification Grammar: modularity and polarization. In *Proceedings of MTT 2005*, Moscou: 163–173.
- Kahane, S. & Mel'čuk, I. A. (1999). Synthèse des phrases à extraction en français contemporain (du réseau sémantique à l'arbre syntaxique). *Traitement Automatique des Langues (T.A.L.)*, 40(2): 25–85.
- Kahane, S. & Polguère, A. (2001). Formal foundations of lexical functions. In *Proceedings of ACL 2001*, Toulouse.
- Kahn, F. (1954). *Le système des temps de l'indicatif (chez un Parisien et chez une Bâloise)*. Droz, Genève.

- Kamp, H. & Rohrer, C. (1983). Tense in texts. In Rainer, B., Schwarze, C. & von Stechow, A. (dir.). *Meaning, use and interpretation of language*. De Gruyter, Berlin: 250–269.
- Kayne, R. (1977). *Syntaxe du français. Le cycle transformationnel*. Seuil, Paris.
- Korzen, H. & Nølke, H. (1990). Projet pour une théorie des emplois du conditionnel. In *Actes du 11^e Congrès des Romanistes Scandinaves*, Trondheim: 273–300.
- Labeau, E. (2002). L'unité de l'imparfait: vues théoriques et perspectives pour les apprenants du français langue étrangère. *Travaux de linguistique*, 45: 157–184.
- Labeau, E. & Larrivée, P. (2005). *Nouveaux développements de l'imparfait*. Rodopi, Amsterdam / New York.
- Lareau, F. (2002). *La synthèse automatique de paraphrases comme outil de vérification des dictionnaires et grammaires de type sens-texte*. Mémoire de maîtrise, Université de Montréal.
- Lareau, F. (2007). Vers une formalisation des décompositions sémantiques dans la Grammaire d'Unification Sens-Texte. In *Actes de TALN 2007 (Vol. 2)*, Toulouse: 163–172.
- Lareau, F. & Wanner, L. (2007). Towards a generic multilingual dependency grammar for text generation. In *Proceedings of the GEAF07 Workshop*, Stanford: 203–223.
- Larousse de la conjugaison*. (1999). Larousse, Paris.
- Laurendeau, P. (2000). L'alternance futur simple / futur périphrastique: une hypothèse modale. *Verbum*, 22(3): 277–292.
- Leeman-Bouix, D. (1994). *Grammaire du verbe français: des formes au sens. Modes, aspects, temps, auxiliaires*. Nathan, Paris.
- Lefevre, F. (1999). *La phrase averbale en français*. L'Harmattan, Paris.
- Le Goffic, P. (dir.) (1986a). *Points de vue sur l'imparfait*. Centre de publications de l'Université de Caen, Caen.

- Le Goffic, P. (1986b). Que l'imparfait n'est pas un temps du passé. In Le Goffic (1986a): 55–69.
- Le Goffic, P. (1997). *Les formes conjuguées du verbe français: oral et écrit*. Ophrys, Paris.
- Martinet, A. (1970). *Éléments de linguistique générale*. Armand Colin, Paris.
- Martinet, A. (1979). *Grammaire fonctionnelle du français*. Crédif / Didier, Paris.
- Mazon, A. (1914). *Emplois des Aspects du Verbe Russe*. Champion, Paris.
- Mel'čuk, I. A. (1973). Towards a linguistic "Meaning-Text" model. In Kiefer, F. (dir.). *Trends in Soviet Theoretical Linguistics*. Reidel, Dordrecht: 33–57.
- Mel'čuk, I. A. (1974). *Opyt teorii lingvističeskix modelej "Smysl ⇔ Tekst"*. *Semantica, Sintaksis*. Nauka, Moskva.
- Mel'čuk, I. A. (1988a). *Dependency Syntax: Theory and Practice*. State University of New York Press, Albany.
- Mel'čuk, I. A. (1988b). Semantic description of lexical units in an explanatory combinatorial dictionary: basic principles and heuristic criteria. *International Journal of Lexicography*, 1: 165–88.
- Mel'čuk, I. A. (1989). Semantic primitives from the viewpoint of the meaning-text linguistic theory. *Quaderni di semantica*, 10(1): 65–102.
- Mel'čuk, I. A. (1993). *Cours de morphologie générale. Introduction et première partie: le mot (Vol. 1)*. Presses de l'Université de Montréal, Montréal / CNRS, Paris.
- Mel'čuk, I. A. (1994). *Cours de morphologie générale. Deuxième partie: significations morphologiques (Vol. 2)*. Presses de l'Université de Montréal, Montréal / CNRS, Paris.
- Mel'čuk, I. A. (1996a). *Cours de morphologie générale. Troisième partie: moyens morphologiques. Quatrième partie: syntactiques morphologiques (Vol. 3)*. Presses de l'Université de Montréal, Montréal / CNRS, Paris.
- Mel'čuk, I. A. (1996b). Lexical functions: A tool for the description of lexical relations in a lexicon. In Wanner, L. (dir.). *Lexical Functions in Lexicography*

- and Natural Language Processing*. John Benjamins, Amsterdam / Philadelphia: 37–102.
- Mel'čuk, I. A. (1997a). *Vers une linguistique Sens-Texte*. Leçon inaugurale au Collège de France, Paris.
- Mel'čuk, I. A. (1997b). *Cours de morphologie générale. Cinquième partie: signes morphologiques (Vol. 4)*. Presses de l'Université de Montréal, Montréal / CNRS, Paris.
- Mel'čuk, I. A. (1997c). Cas grammaticaux, construction verbale de base et voix en massaï: vers une meilleure analyse des concepts. *Bulletin de la Société de linguistique de Paris*, 92(1): 49–110.
- Mel'čuk, I. A. (2000). *Cours de morphologie générale. Sixième partie: modèles morphologiques. Septième partie: principes de la description morphologique (Vol. 5)*. Presses de l'Université de Montréal, Montréal / CNRS, Paris.
- Mel'čuk, I. A. (2001a). Language: Dependency. In Smelser, N. J. & Baltes, P. B. (dir.) *International Encyclopedia of the Social and Behavioral Sciences*. Elsevier Science, Oxford: 8336–8344.
- Mel'čuk, I. A. (2001b). Grammatical voice in French (a short description in the Meaning-Text framework). In Chrakovskij *et al.* (dir.). *Studies on the Syntax and Semantics of Slavonic Languages (Papers in Honour of Andrzej Bogusławski on the Occasion of his 70th Birthday)*. Universität Oldenburg: 265–290.
- Mel'čuk, I. A. (2001c). *Communicative organization in natural language: the semantic-communicative structure of sentences*. Benjamins, Amsterdam / Philadelphia.
- Mel'čuk, I. A. (2004). Actants in Semantics and Syntax. *Linguistics*, 42(1): 1–66 (première partie) et 42(2): 247–291 (deuxième partie).
- Mel'čuk, I. A. (2006) Explanatory Combinatorial Dictionary. In Sica, G. (dir.). *Open Problems in linguistics and lexicography*. Polimetrica, Monza: 225–355.

- Mel'čuk, I. A. *et al.* (1984–1999). *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques*, vol. I à IV, Les Presses de l'Université de Montréal, Montréal.
- Mel'čuk, I. A., Clas, A. & Polguère, A. (1995). *Introduction à la lexicologie explicative et combinatoire*. Duculot, Louvain-la-Neuve.
- Mel'čuk, I. A. & Pertsov, N. (1987). *Surface Syntax of English: A Formal Model Within the Meaning-Text Framework*. John Benjamins, Amsterdam.
- Mel'čuk, I. A. & Polguère, A. (2007). *Lexique actif du français. L'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. De Boeck & Larcier, Bruxelles.
- Mel'čuk, I. A. & Polguère, A. (dir.). (2008). *Dependency in Linguistic Description*. John Benjamins, Amsterdam / Philadelphia.
- Menanteau, D. (1986). Le mode verbal, classe grammaticale? *La linguistique*, 22(1): 69–80.
- Milićević, J. (1997). *Étiquettes sémantiques dans un dictionnaire formalisé du type Dictionnaire Explicatif et Combinatoire*. Mémoire de maîtrise, Université de Montréal.
- Milićević, J. (2006). A short guide to the Meaning-Text theory. *Journal of Koralex*, 8: 187–233.
- Milićević, J. (2007). *La paraphrase: Modélisation de la paraphrase langagière*. Peter Lang, Bern.
- Molendijk, A. (1990). *Le passé simple et l'imparfait: une approche reichenbachienne*. Rodopi, Amsterdam / Atlanta.
- Mourelatos, A. P. (1978). Events, processes and states. *Linguistics and Philosophy*, 2: 415–434.
- Morin, Y.-C. (1987). Remarques sur l'organisation de la flexion des verbes français. *International Review of Applied Linguistics*, 77–78: 13–91.

- Morin, Y.-C. (2001). La troncation des radicaux verbaux en français depuis le moyen âge: Études diachroniques. *Recherches linguistiques de Vincennes*, 30: 63–86.
- Nasr, A. (1995). A formalism and a parser for lexicalized dependency grammars. In *4th International Workshop on Parsing Technologies*, Prague: 186–195.
- Partee, B. H. (1973). Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy*, 70: 601–609.
- Partee, B. H. (1984). Nominal and temporal anaphora. *Linguistics and Philosophy*, 7: 243–286.
- Perrier, G. (2002). Descriptions d'arbres avec polarités: les grammaires d'interaction. In *Actes de TALN 2002*, Nancy.
- 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.
- Polguère, A. (1998a). La théorie Sens-Texte. *Dialangue*, 8–9: 9–30.
- Polguère, A. (1998b). Pour un modèle stratifié de la lexicalisation en génération de texte. *TAL*, 39(2): 57–76.
- Polguère, A. (2000). Une base de données lexicales du français et ses applications possibles en didactique. *LIDIL*, 21: 75–97.
- Polguère, A. (2003a). Étiquetage sémantique des lexies dans la base de données DiCo. *Traitement Automatique des Langues (T.A.L.)*, 44(2): 39–68.
- Polguère, A. (2003b). *Lexicologie et sémantique lexicale. Notions fondamentales*. Presses de l'Université de Montréal, Montréal.
- Polguère, A. (2008). *Lexicologie et sémantique lexicale. Notions fondamentales* (2^e édition). Presses de l'Université de Montréal, Montréal.
- Pollard, C. J. & Sag, I. A. (1994). *Head-driven phrase structure grammar*. Center for the Study of Language and Information, Stanford / University of Chicago Press, Chicago.

- Psyché, V. (2003). *État de l'art sur l'ontologie et application de l'ingénierie ontologique au domaine du téléapprentissage*. Rapport de recherche, Centre de recherche LICEF, Téléuniversité de l'Université du Québec, Montréal.
- Reichenbach, H. (1947). *Elements of symbolic logic*. Macmillan, New York.
- Rey, A. & Rey-Debove, J. (dir.) (1986). *Le Petit Robert 1*. Le Robert, Paris.
- Riegel, M., Pellat, J.-C. & Rioul, R. (1994). *Grammaire méthodique du français*. Presses universitaires de France, Paris.
- Sapir, E. (1915). Abnormal types of Speech in Nootka. *Canada, Geological Survey, Memoir 62, Anthropological Series, 5*. [Réimprimé dans: Sapir, E. (1973). *Selected Writings of Edward Sapir in Language, Culture and Personality*. University of California Press, Berkeley].
- Tesnière, L. (1939). Théorie structurale des temps surcomposés. In *Mélanges de linguistique offerts à Charles Bally*. Georg, Genève.
- Tesnière, L. (1959). *Éléments de syntaxe structurale*. Klincksieck, Paris.
- Te Winkel, L. A. (1866). Over de wijzen en tijden der werkwoorden. *De Taalgids*, 8: 66–75.
- Touratier, C. (1996). *Le système verbal français*. Armand Colin, Paris.
- Van Den Eynde, K. & Blanche-Benveniste, C. (1970). Essai d'analyse morphologique du verbe français. *Orbis*, 19: 404–429.
- Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 66: 143–160.
- Verkuyl, H. (2005). How (in)sensitive is tense to aspectual information? In Hollebrandse, B., van Hout, A. & Vet, C. (dir.). *Crosslinguistic Views on Tense, Aspect and Modality*. Rodopi, Amsterdam / New York: 145–169.
- Vet, C. (1980). *Temps, aspects et adverbes de temps en français contemporain. Essai de sémantique formelle*. Droz, Genève.
- Vetters, C. (1996). *Temps, aspect et narration*. Rodopi, Amsterdam.
- Wagner, R. L. & Pinchon, J. (1991 [1962]). *Grammaire du français classique et moderne*. Hachette, Paris.

- Wanner, L. (1992). Lexical choice and the organization of lexical resources in text generation. In *Proceedings of the 10th European conference on Artificial intelligence*, Vienne: 495–499.
- Wanner, L. (1997). *Exploring Lexical Resources for Text Generation in a Systemic Functional Language Model*. Thèse de doctorat, Universität des Saarlandes, Saarbrücken.
- Wanner, L., Nicklaß, D., Bohnet, B., Bouayad-Agha, N., Bronder, J., Ferreira, F., Friedrich, F., Karppinen, A., Lareau, F., Lohmeyer, A., Panighi, A., Parisio, S., Scheu-Hachtel, H. & Serpa, J. (2007). From measurement data to environmental information: Marquis — a multimodal air quality information service for the general public. In *Proceedings of ISESS 2007*, Prague.
- Weinrich, H. (1973). *Le temps*. Le Seuil, Paris.
- Wierzbicka A. (1996). *Semantics: primes and universals*. Oxford University Press, Oxford / Toronto.
- Wilmet, M. (2003). *Grammaire critique du français* (3^e éd.). Duculot, Bruxelles.
- Yoo, Y. (2003). *Ordre des mots et prosodie: essai de description et de formalisation pour le grec moderne*. Thèse de doctorat, Université Paris 7.
- Yoo, Y. & Gerdes, K. (2004). A dependency account of Korean Word Order. In *Proceedings of the Linguistic Society of Korea 2004*, Séoul.
- Yvon, H. (1951). Convient-il de distinguer dans le verbe français des temps relatifs et des temps absolus? *Le français moderne*, 19: 265–276.
- Žolkovskij, A. & Mel'čuk, I. A. (1965). O vozmožnom metode i instrumentax semantičeskogo sinteza. *NTI*, 5: 23–28.
- Žolkovskij, A. & Mel'čuk, I. A. (1967). O semantičeskom sinteze. *Problemy kibernetiki*, 19: 177–238.
- Žolkovskij, A. & Mel'čuk, I. A. (1970). Sur la synthèse sémantique. *T. A. Informations*, 2: 1–85.

Annexe A. Les formes de PRENDRE

Voici une liste des formes synthétiques et analytiques du verbe **PRENDRE** (ou du verbe **PARTIR**, pour illustrer l'utilisation de l'auxiliaire **ÊTRE**) que nous considérons dans le cadre de cette thèse. Pour les formes finies, nous faisons abstraction ici de la voix, de la personne et du nombre (nous n'utilisons que les formes actives à la troisième personne du singulier, sauf pour l'impératif qui apparaît à la seconde personne du singulier). Pour les participes, nous ne tenons pas compte de l'accord en genre et en nombre.

1 Les formes de base

1. *prend*
2. *prenait*
3. *prendra*
4. *prit*
5. *prendrait*
6. *prenne*
7. *prends!*
8. *prendre*
9. *prenant*
10. *pris*
11. *en prenant*

2 Les formes composées

2.1 AVOIR <ÊTRE>

12. *a pris <est parti>*
13. *avait pris <était parti>*
14. *aura pris <sera parti>*
15. *eut pris <fut parti>*
16. *aurait pris <serait parti>*
17. *ait pris <soit parti>*
18. *aie pris <sois parti>*
19. *avoir pris <être parti>*
20. *ayant pris <étant parti>*
21. *en ayant pris <en étant parti>*

2.2 ALLER

- 22. *va prendre*
- 23. *allait prendre*

2.3 VENIR

- 24. *vient de prendre*
- 25. *venait de prendre*
- 26. *viendra de prendre*
- 27. *vint de prendre*
- 28. *viendrait de prendre*
- 29. *vienne de prendre*
- 30. *venir de prendre*
- 31. *venant de prendre*
- 32. *en venant de prendre*

2.4 「EN TRAIN」

- 33. *est en train de prendre*
- 34. *était en train de prendre*
- 35. *sera en train de prendre*
- 36. *fut en train de prendre*
- 37. *serait en train de prendre*
- 38. *soit en train de prendre*
- 39. *être en train de prendre*
- 40. *étant en train de prendre*

2.5 「SUR LE POINT」

- 41. *est sur le point de prendre*
- 42. *était sur le point de prendre*
- 43. *sera sur le point de prendre*
- 44. *fut sur le point de prendre*
- 45. *serait sur le point de prendre*
- 46. *soit sur le point de prendre*
- 47. *être sur le point de prendre*
- 48. *étant sur le point de prendre*

3 Les formes doublement composées

3.1 AVOIR + AVOIR <ÊTRE>

- 49. *a eu pris <a été parti>*
- 50. *avait eu pris <avait été parti>*
- 51. *aurait eu pris <aurait été parti>*
- 52. *aura eu pris <aura été parti>*

3.2 AVOIR + 「EN TRAIN」

- 53. *aurait été en train de prendre*

3.3 AVOIR + 「SUR LE POINT」

- 54. *aurait été sur le point de prendre*

3.4 ALLER + AVOIR <ÊTRE>

- 55. *va avoir pris <va être parti>*
- 56. *allait avoir pris <allait être parti>*

3.5 ALLER + VENIR

- 57. *va venir de prendre*
- 58. *allait venir prendre*

3.6 ALLER + 「EN TRAIN」

- 59. *va être en train de prendre*
- 60. *allait être en train de prendre*

3.7 ALLER + 「SUR LE POINT」

- 61. *va être sur le point de prendre*
- 62. *allait être sur le point de prendre*

3.8 「SUR LE POINT」 + AVOIR <ÊTRE>

- 63. *? est sur le point d'avoir pris <? d'être parti>*
- 64. *? était sur le point d'avoir pris <? d'être parti>*
- 65. *? sera sur le point d'avoir pris <? d'être parti>*

66. ?serait sur le point d'avoir pris <? d'être parti>
 67. ?fut sur le point d'avoir pris <? d'être parti>
 68. ?soit sur le point d'avoir pris <? d'être parti>
 69. ?être sur le point d'avoir pris <? d'être parti>
 70. ?étant sur le point d'avoir pris <? d'être parti>

4 Les formes triplement composées

4.1 AVOIR + 「SUR LE POINT」 + AVOIR / ÊTRE

71. ?aurait été sur le point d'avoir pris <? d'être parti>

4.2 ALLER + 「SUR LE POINT」 + AVOIR <ÊTRE>

72. ?va être sur le point d'avoir pris <? d'être parti>
 73. ?allait être sur le point d'avoir pris <? d'être parti>

Annexe B. L'encodage XML des grammaires GUST

Il est crucial de fixer un format d'échange pour les grammaires si on souhaite les voir utilisées dans des systèmes de traitement automatique des langues. Nous proposons donc ici un encodage XML pour les grammaires GUST.

Les grammaires sont contenues à l'intérieur de balises `<grammar>...</grammar>`. On identifie la langue décrite par la grammaire en question par l'attribut `language`, qui prend comme valeur le code de trois lettres correspondant à la langue décrite selon le standard ISO 639¹. Par exemple, pour la grammaire du français :

```
<grammar language=FRE>
...
</grammar>
```

Chaque grammaire est constituée de modules. Les modules sont contenus à l'intérieur des balises `<module>...</module>`, dont l'attribut `name` spécifie le nom du module en question. Par exemple, pour la grammaire de bonne formation sémantique :

```
<module name=Gsem>
...
</module>
```

Les modules contiennent un nombre illimité de règles. Chaque règle est représentée à l'intérieur des balises `<structure>...</structure>`. L'attribut `name` est également utilisé pour attribuer un nom à la règle en question. Par exemple, nous appellerons «manger» la règle qui décrit la valence du sémantème 'manger'. Nous utilisons également un attribut `id` pour associer un numéro d'identification unique à chaque structure. Il est alors possible de faire référence à une structure en particulier (par exemple, dans un système de génération de texte, on peut vouloir voir la

1. Ce standard permet pour certaines langues deux codes : un basé sur le nom de la langue en anglais et l'autre sur son nom dans la langue elle-même. Par exemple, pour le français, il existe les codes «FRE» et «FRA». Nous choisirons toujours le code basé sur le nom anglais.

liste de toutes les structures qui ont été combinées pour générer une phrase donnée).

```
<structure id=1 name=manger>
...
</structure>
```

Un module peut également contenir une ou plusieurs structure(s) initiale(s). Ces structures sont décrites comme toutes les autres règles, à l'exception près qu'elles possèdent un attribut `initial=yes` dans la balise `<structure>`. La balise `<module>` doit également avoir un attribut `initial` dont la valeur est la liste des numéros d'identification des structures initiales du module. Par exemple, si dans le module syntaxique les règles 1 et 2 sont des structures initiales, on aura :

```
<module name=Gsynt initial="1,2">
  <structure id=1 name="sommet verbe fini" initial=yes>
    ...
  </structure>
  <structure id=2 name="sommet interjection" initial=yes>
    ...
  </structure>
  ...
</module>
```

À l'intérieur des balises `<structure>...</structure>`, on trouvera quatre éléments. D'abord, des commentaires qui expliquent le phénomène linguistique que la règle en question tente de modéliser. Ces commentaires ne doivent pas être inscrits comme des commentaires XML, qui seraient ignorés par un parseur. Ils doivent être contenus à l'intérieur d'une balise `<comments>...</comments>`. Ainsi, ils peuvent être récupérés par un programme pour être présentés à l'utilisateur si nécessaire. Afin d'illustrer le phénomène décrit par la règle, on peut également ajouter des exemples. On prévoit à cette fin une balise `<examples>...</examples>` qui contiendra un certain nombre de phrases, chacune contenue à l'intérieur d'une balise `<example>...</example>`. Les éléments `<comments>...</comments>` et `<examples>...</examples>` sont facultatifs. Par contre, toute structure doit être constituée d'un ensemble d'objets, contenus à l'intérieur des balises `<objects>...</objects>`, et d'un ensemble de fonctions contenues dans les balises `<functions>...</functions>`.

```

<structure id=1 name=manger>
  <comments>
    'manger' a deux actants sémantiques
  </comments>
  <examples>
    <example>
      Pierre mange du gâteau
    </example>
  </examples>
  <objects>
    ...
  </objects>
  <functions>
    ...
  </functions>
</structure>

```

Chaque objet est encodé par une balise simple `<object/>` qui a deux attributs obligatoires. L'attribut `id` prend comme valeur un numéro, différent pour chaque objet de la structure, qui sert d'identificateur pour l'objet en question alors que l'attribut `type` donne, comme son nom l'indique, son type. La valeur de l'attribut `type` ne peut être qu'un des types définis pour la grammaire en question.

Chaque fonction est encodée par une balise simple `<function/>` qui doit avoir quatre attributs. Il s'agit des attributs `name`, `type`, `argument` et `value`, qui indiquent respectivement le nom de la fonction et son type, l'objet auquel elle s'applique et la valeur qu'elle retourne. L'attribut `type` ne peut prendre comme valeur que `polarity`, `structural` ou `label` (il s'agit des trois principaux types de fonctions identifiés au Chapitre VII, § 2, p. 226). Enfin, la valeur de l'attribut `value` peut être une chaîne de caractères quelconque (dans le cas des fonctions d'étiquetage), le nom d'une des polarités (pour les fonctions de polarisation) ou encore le numéro d'identification d'un objet de la même structure (pour les fonctions structurales). La structure XML suivante encode la règle sémantique $\langle \{ 1:\text{nœud-sém}, 2:\text{nœud-sém}, 3:\text{nœud-sém}, 4:\text{arc-sém}, 5:\text{arc-sém} \}, \{ \text{étiquette}(1)=\text{"manger"}, \text{étiquette}(4)=\text{"1"}, \text{étiquette}(5)=\text{"2"}, \text{source}(4)=1, \text{cible}(4)=2, \text{source}(5)=1, \text{cible}(5)=3, p_{sém}(1)=\blacksquare, p_{sém}(2)=\square, p_{sém}(3)=\square, p_{sém}(4)=\blacksquare, p_{sém}(5)=\blacksquare} \} \rangle$.

Enfin, chaque système qui utilise des ressources GUST pourra ajouter à certaines balises des attributs qui lui sont propres pour stocker certaines informations.

```

<structure id=1 name=manger>
  <comments>
    'manger' a deux actants sémantiques
  </comments>
  <examples>
    <example>
      Pierre mange du gâteau
    </example>
  </examples>
  <objects>
    <object id=1 type=semnode/>
    <object id=2 type=semnode/>
    <object id=3 type=semnode/>
    <object id=4 type=semedge/>
    <object id=5 type=semedge/>
  </objects>
  <functions>
    <function name=label type=label argument=1 value="manger"/>
    <function name=label type=label argument=4 value="1"/>
    <function name=label type=label argument=5 value="2"/>
    <function name=source type=structural argument=4 value=1/>
    <function name=target type=structural argument=4 value=2/>
    <function name=source type=structural argument=5 value=1/>
    <function name=target type=structural argument=5 value=3/>
    <function name=psem type=polarity argument=1 value=black/>
    <function name=psem type=polarity argument=2 value=white/>
    <function name=psem type=polarity argument=3 value=white/>
    <function name=psem type=polarity argument=4 value=black/>
    <function name=psem type=polarity argument=5 value=black/>
  </functions>
</structure>

```

Par exemple, dans un outil d'édition de grammaires, il peut être utile d'ajouter à chaque balise une date de modification ou un numéro de version, de façon à faciliter la coordination du travail lorsque plusieurs linguistes travaillent sur les mêmes ressources.