



HAL
open science

Un modèle dynamique et parcimonieux du traitement automatisé de l'aspect dans les langues naturelles

Damien Munch

► **To cite this version:**

Damien Munch. Un modèle dynamique et parcimonieux du traitement automatisé de l'aspect dans les langues naturelles. Linguistique. Télécom ParisTech, 2013. Français. NNT : 2013ENST0058 . tel-01226505

HAL Id: tel-01226505

<https://pastel.hal.science/tel-01226505>

Submitted on 9 Nov 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

EDITE ED 130

Doctorat ParisTech

THÈSE

pour obtenir le grade de docteur délivré par

Télécom ParisTech

Spécialité “ Informatique et réseaux ”

présentée et soutenue publiquement par

Damien MUNCH

le 5 Novembre 2013

**Un modèle dynamique et parcimonieux
du traitement automatisé de l'aspect dans les langues naturelles**

Directeur de thèse : Jean-Louis DESSALLES

Jury

M. Elie NAJM, Professeur, Télécom ParisTech

Mme Delphine BATTISTELLI, Professeure, Université Paris Ouest Nanterre La Défense

M. Patrice ENJALBERT, Professeur, Université de Caen

Mme. Ada DIACONESCU, Maître de conférence, Télécom ParisTech

M. Luc STEELS Professeur, Vrije Universiteit Brussel

M. Jean-Louis DESSALLES, Maître de conférence, Télécom ParisTech

Président

Rapporteure

Rapporteur

Examinatrice

Examineur

Directeur de thèse

Télécom ParisTech

école de l'Institut Mines Télécom – membre de ParisTech

46, rue Barrault – 75634 Paris Cedex 13 – Tél. + 33 (0)1 45 81 77 77 – www.telecom-paristech.fr

Un modèle dynamique et parcimonieux du traitement automatisé de l'aspect dans les langues naturelles

RÉSUMÉ : Dans cette thèse nous avons cherché et développé un modèle du traitement de l'aspect dans les langues naturelles. Notre objectif a été d'élaborer un modèle détaillé et explicatif qui montre la possibilité de traiter l'aspect sur un nombre choisi d'énoncés tout en suivant des contraintes fortes de parcimonie et de plausibilité cognitive. Nous avons réussi à mettre au point un modèle original dans sa réalisation, mais aussi dans ses résultats : des explications nouvelles sont données pour le traitement d'interprétations comme la répétition, la perfectivité ou l'inchoativité ; tout en dévoilant un phénomène original dit de "prédication".

Mots clés : Temporalité, Temps, Aspect, Traitement Automatique du Langage, Parcimonie, Minimalisme, Plausibilité cognitive

A Dynamic and Parsimonious model for the Processing of Aspect in Natural Language

ABSTRACT : The purpose of this work is to design and to implement a computational model for the processing of aspect in natural language.

Our goal is to elaborate a detailed and explicative model of aspect. This model should be able to process aspect on a chosen number of sentences, while following strong constraints of parsimony and cognitive plausibility.

We were successful in creating such a model, with both an original design and an extensive explanatory power. New explanations have been obtained for phenomena like repetition, perfectivity and inchoativity. We also propose a new mechanism based on the notion of "predication".

Keywords : Time, Tense, Aspect, Natural Language Processing, Parsimony, Minimalism, Cognitive plausibility

Merci à Jean-Louis, qui a encadré ces travaux mieux que je n'aurais pu l'imaginer. Merci à toute l'équipe de DBWeb et mes collègues d'INFRES pour tout les échanges, studieux ou ludiques, qui ont permis d'éclairer mes années de thèses. Merci à mes parents et ma famille, qui malgré la distance ont su m'encourager à chaque instants. Merci enfin à Marilena, pour tout nos échanges, ses conseils avisés, et pour m'avoir supporté pendant tout ce temps.

Damien Munch

Sommaire

1 Introduction.....	10
1.1 Le compréhension de la langue.....	10
1.2 Problématique.....	11
1.2.1 Temps.....	11
1.2.2 Modalité.....	12
1.2.3 Aspect.....	13
1.2.4 Modélisation.....	14
1.3 Contraintes.....	15
1.4 Modèles existants sur le temps et l'aspect.....	16
1.5 Points clés de notre approche.....	17
1.6 Choix des énoncés.....	18
2 Contraintes.....	19
2.1 La parcimonie.....	19
2.1.1 Le modèle scientifique.....	19
2.1.2 La complexité de Kolmogorov.....	21
2.2 La plausibilité cognitive.....	26
3 Approches non linguistiques du temps.....	32
3.1 En philosophie.....	32
3.2 En logique.....	33
3.2.1 Application au langage naturel.....	39
3.3 En linguistique cognitive.....	40
4 Approches linguistique du temps.....	42
4.1 L'aspect.....	42
4.1.1 Classes aspectuelles.....	42
4.1.2 Durativité.....	43
4.1.3 Ordre de grandeur.....	44
4.1.4 Dynamicité.....	45
4.1.5 Perfectivité et perspective.....	46
4.1.6 Télicité.....	48

4.1.7 Point d'achèvement.....	50
4.1.8 Homogénéité.....	52
4.1.9 Fréquence.....	53
4.1.10 Phasage.....	54
4.2 Temporalité, modalité et autres éléments pour le traitement du temps.....	55
4.2.1 Passé-Présent-Futur.....	55
4.2.2 Moment de référence.....	58
4.2.3 Opérateurs relationnels.....	60
4.2.4 Détermination.....	61
4.3 Modalité.....	62
4.4 Attitude propositionnelle.....	63
4.5 Pragmatique.....	65
5 Des modélisations pour le traitement de l'aspect.....	67
5.1 Reichenbach.....	67
5.2 Gosselin.....	71
5.3 Kamp et Reyle.....	77
5.4 De Swart.....	82
6 Élaboration du modèle.....	86
6.1 Une approche modulaire.....	86
6.2 La fusion simultanée dans la syntaxe et la sémantique.....	90
6.3 Le module perceptif.....	94
6.3.1 Le système conceptuel.....	94
6.3.2 L'image perceptive.....	97
6.3.2.1 La fusion perceptive.....	98
6.3.2.2 Images et informations sémantiques.....	99
6.3.2.3 La durée.....	101
6.3.2.4 Le type aspectuel.....	104
6.3.3 L'ordonnement.....	107
6.3.3.1 Connaissances et causalité.....	108
6.3.3.2 Des images relationnelles.....	109
6.3.3.3 Un calcul de l'ordonnement.....	112
6.3.4 Résumé des propriétés du module perceptif.....	113

7	Description du modèle.....	115
7.1	Vue d'ensemble.....	115
7.2	AIS : La Structure d'Information Aspectuelle.....	117
7.2.1	Motivations.....	117
7.2.2	Une variable binaire dynamique : Le switch.....	120
7.2.3	ImageID.....	120
7.2.4	Viewpoint.....	121
7.2.5	Détermination temporelle.....	123
7.2.6	Durativité.....	125
7.2.7	Multiplicity.....	126
7.2.8	Les AIS dans le modèle.....	126
7.2.9	Détails d'implémentation.....	128
7.3	aMerge : La fusion sémantique (aspectuelle).....	129
7.3.1	Interactions du module aspectuel.....	131
7.3.2	Unification.....	133
7.3.2.1	Unifications avec contraintes perceptives.....	133
7.3.2.2	Unification sans contraintes.....	136
7.3.3	Opérations.....	136
7.3.3.1	Focus-in.....	137
7.3.3.2	Focus-out et "prédication" temporelle.....	138
7.3.3.3	Repetition.....	139
7.3.3.4	Slice.....	142
7.3.3.5	Shift.....	144
7.4	Lecture des arbres de AIS.....	148
8	Traitements.....	151
8.1	Exemples de traitements.....	151
8.1.1	Un cas détaillé.....	151
8.1.2	Exemple commenté.....	157
8.2	Implémentation.....	158
8.2.1	Précisions techniques.....	158
8.2.2	Résultats.....	160
9	Calcul de la parcimonie.....	161

9.1	Liste d'énoncés.....	162
9.2	Complexité du null-model.....	164
9.3	Complexité du modèle avec AIS et aMerge.....	166
9.3.1	Les forces du modèle.....	166
9.3.2	Occupation des AIS et informations perceptuelles.....	168
9.3.3	Calcul de la complexité de l'algorithme.....	168
9.4	Conclusion sur la parcimonie.....	169
10	Limites et Perspectives.....	171
10.1	Production et interprétation.....	171
10.2	la temporalité et l'aspect.....	171
10.3	Le module perceptif.....	173
10.4	Les dimensions sémantiques.....	174
11	Conclusion.....	175
11.1	Apports de la thèse.....	175
11.2	Avenir.....	177
12	Index.....	179
13	Bibliographie.....	181
14	Annexes.....	187
14.1	Listes d'exemples et de leurs interprétations.....	188
14.1.1	Exemples avec l'imparfait.....	188
14.1.2	Exemples avec le passé composé.....	190
14.1.3	Avec apposition.....	192
14.2	Extraits de l'implémentation.....	194
14.2.1	syn_file.pl.....	194
14.2.2	per_file.pl.....	194
14.2.3	sem_file.pl.....	195
14.2.4	tmt_perception.pl.....	197
14.2.5	tmt_model.pl.....	197

1 Introduction

1.1 Le compréhension de la langue

"Il était une fois ...". Quatre mots qui nous font voyager dans un autre monde ou à une époque différente. L'expression est connue, elle est ancrée dans notre mémoire depuis tout petit, et quand nous la lisons ou l'entendons, nous sommes immédiatement projetés dans le passé, celui de notre enfance ou celui de l'histoire où se sont déroulés bien des aventures.

"Quand je serai grand...". Ce début d'histoire aussi nous la connaissons bien ; elle nous renvoie à notre enfance, mais aussi et surtout au futur de notre enfance. Aujourd'hui. Demain. Ce sont des mots d'imagination, de projection dans un avenir.

Ces deux débuts de phrases ne sont pas seulement des expressions toutes faites, mais des constructions de la langue. Ils suscitent chez ceux qui les lisent bien des images, même si nous les entendons pour la première fois. Comment ? Comment l'être humain est-il capable d'imaginer, de construire un sens, uniquement en observant une liste de symboles ? Plus précisément, nous comprenons que la première expression renvoie à un passé, certain ou hypothétique, et l'autre à un futur possible.

Ces informations temporelles, données par les multiples lemmes et grammèmes, ont dû être comprises par le lecteur et traitées sans efforts ou calculs conscients.

Mais toutes les phrases sont-elles aussi faciles à comprendre ? Et si les humains, y compris les tous jeunes enfants, y arrivent, un programme informatique pourrait-il en être capable ?

1.2 Problématique

Le traitement du langage, qu'il soit réalisé par un être humain ou un logiciel informatique, va se heurter à une multitude de problèmes. Pour faciliter leurs résolutions, le langage a été divisé en différentes parties plus ou moins interconnectés : L'une des catégorisations linguistiques les plus emblématiques du langage est la partition entre **syntaxe**, **sémantique**, **pragmatique** et **phonétique**. Une autre division, transversale à celle que nous venons de donner, est de distinguer les différentes catégories de sens fournies : temps, détermination, genre, nom, etc.

Nous nous sommes à l'origine concentré sur la **temporalité** dans la langue, mais nous avons constaté qu'il était nécessaire de découper notre sujet une nouvelle fois pour espérer en traiter correctement : **temps**, **modalité** et **aspect**. Cette division est commune en linguistique et permet de catégoriser l'essentiel des temps grammaticaux des langues occidentales. La distinction est restée, même s'il est apparu que chacune est intrinsèquement liée aux autres et qu'il est difficile de se limiter à une seule.

Voyons rapidement en quoi consistent ces trois catégories avant de préciser ce que nous entendons par "traitement du langage".

1.2.1 Temps

Le temps décrit les relations entre des situations ou/et des "moments particuliers". C'est le temps qui permet d'exprimer si une situation est passée, présente ou future, en se positionnant par rapport à un "moment du présent", que l'on décrit parfois comme le moment d'énonciation ou le "maintenant". Le temps est typiquement exprimé en français par les temps de conjugaison. Le passé composé exprimera un passé ; le présent ... le présent ou un futur proche ; et le futur simple un futur.

- (1) *Jean a couru le marathon.*
- (2) *Jean court le marathon.*
- (3) *Jean courra le marathon.*

Le temps peut aussi être obtenu grâce à d'autres éléments du lexique, comme des adverbes : *hier* est obligatoirement le jour passé précédant le jour présent ; ou des syntagmes nominaux : *la veille*

se contente de signifier une précedence par rapport à un jour de référence (qui peut très bien prendre place dans le futur).

(4) *La veille du concours, elle ira ...*

(5) *# Hier, elle ira ...*¹

Sans référence au "moment du présent", il est tout à fait possible de positionner les situations par rapport à elles-mêmes ou à d'autres "moments" de référence.

(6) *Elle était malade le jour avant son anniversaire.*

Le positionnement n'est pas forcément explicite et peut très bien être déduit à partir de connaissances générales :

(7) *Marie est tombé. Jean l'a poussé.*

On déduit dans (7) que Jean a poussé Marie, et qu'ensuite celle-ci est tombée. Cet enchaînement est privilégié en raison du lien de causalité entre pousser et tomber. Il est néanmoins possible d'interpréter (7) différemment et de comprendre que Jean pousse Marie après qu'elle soit tombée.

1.2.2 Modalité

La modalité est une information qui peut être portée par les temps de conjugaison, ce qui explique pourquoi elle est associée à la temporalité. Elle décrit les notions de chances de réalisation, d'obligation, d'appréciation. En français, il est possible d'exprimer différentes modalités en fonction du temps de conjugaison et particulièrement le conditionnel.

(8) *Demain, Marie parlera.*

(9) *Demain, Marie va parler.*

(10) *Demain, Marie parlerait [de ce sujet].*

On peut exprimer la modalité en précisant le point de vue d'un sujet :

(11) *Jean prétend qu'elle va parler.*

1 Pour les personnes non familières des notations linguistiques, un astérisque * en début d'énoncé signifie que celui-ci est considéré comme grammaticalement incorrect. Si l'énoncé est précédé d'un ou plusieurs points d'interrogations ?, il sera considéré comme "plus ou moins" correct. Un symbole dièse # exprimera un énoncé grammaticalement correct mais dont l'interprétation est difficile, ou "bizarre".

(12) *Jean suppose qu'elle va parler.*

Ou utiliser des adverbes :

(13) *Elle est sûrement partie.*

(14) *Elle est peut-être partie.*

Certaines interprétations peuvent être particulièrement ambiguës :

(15) *[...] Deux secondes plus tard, et le train déraillait.*

Pour (15) on peut comprendre que, dans le passé, le train a effectivement déraillé deux secondes après un événement ; ou que cet événement n'a pas eu lieu, mais que si cela avait été le cas, le train aurait très certainement déraillé.

1.2.3 Aspect

L'aspect est la description du développement d'une situation (on pourra parler pour simplifier d'un événement ou d'une action dans leurs acceptions communes) : commencement, déroulement, achèvement, répétition, etc.

Il n'existe pas de définition communément acceptée de l'aspect, même si tout le monde s'accorde pour reconnaître la notion d'aspect.

En français, la distinction entre imparfait et passé simple (ou passé composé pour la langue parlée) est d'abord aspectuelle :

(16) *Marie était malade la semaine dernière.* (et l'est peut-être encore. "Imperfectif")

(17) *Marie a été malade la semaine dernière.* (et ne l'est plus. "Perfectif")

Des effets de répétitions ou de commencements peuvent être compris alors même qu'aucun élément du lexique n'en fait explicitement part :

(18) *Marie a éternué pendant le spectacle.* (Répétition possible des éternuements, ou un unique éternuement)

(19) *Marie éternuait pendant le spectacle.* (Répétition obligatoire)

(20) *Jean a gravi le sommet en deux heures.* (*graver le sommet a duré deux heures*)

(21) *Jean a atteint le sommet en deux heures.* (*atteindre le sommet n'a pas duré deux heures,*

mais a eu lieu après deux heures)

Les énoncés (20) et (21) sont très semblables, mais leur interprétation : Dans (20) l'ascension aura duré exactement deux heures, alors que dans (21) l'ascension (qui n'en est pas forcément une, Jean a très bien pu prendre un hélicoptère pour *atteindre le sommet*) a pu commencer bien avant les deux heures indiqués au moment de l'énonciation de la phrase.

L'aspect peut changer avec la situation, sans changer la conjugaison :

(22) *Jean dessine un cercle.* (et le termine. *Dessiner un cercle* est "télique")

(23) *Jean dessine (des natures mortes).* (c'est un "état")

On distingue généralement deux types d'aspect :

L'aspect dit "grammatical", qui peut être **marqué** par des lexèmes ou des grammèmes. Ce sont les phénomènes que l'on obtient avec les énoncés (16) à (21) : perfectivité, répétition, inchoativité, etc.

Et l'aspect dit "lexical" (aussi appelé *Aktionsart*, terme que nous préférons par la suite), qui serait une propriété inhérente à la situation décrite par un verbe ou un syntagme verbal. C'est le cas avec les exemples (22) et (23).

C'est sur l'aspect, et plus particulièrement l'aspect "grammatical" que nous nous sommes concentré pendant nos recherches, en essayant autant que possible de ne pas faire appel au temps et à la modalité dans nos explications.

Nous parlerons d'**aktionsart** lorsque nous ferons référence à l'aspect "lexical", et d'**aspect** tout court pour l'aspect "grammatical".

1.2.4 Modélisation

Dans cette thèse nous avons cherché et développé un **modèle du traitement de l'aspect**.

Notre objectif n'a pas été de décrire et catégoriser – comme en Linguistique – ni de réaliser un système informatique capable de traiter un maximum d'énoncés – comme en Traitement Automatique du Langage (TAL) – mais de disposer d'un modèle détaillé qui montre qu'**il est possible de traiter l'aspect** sur un nombre choisi d'énoncés tout **en suivant un certains nombre de contraintes de parcimonie et de plausibilité cognitive**. Notre espoir est de voir

ces contraintes conduire à des améliorations des modèles du traitement du langage, quitte à réviser plus tard certaines des hypothèses faites. On espère ainsi fournir une nouvelle mesure pour évaluer **comment** les énoncés sont traités : La parcimonie.

Si nous nous limitons à traiter d'aspect, nous n'avons pas voulu nous contraindre à un seul des ensembles que sont la syntaxe, la sémantique ou la pragmatique. Notre modèle se situe au niveau de la sémantique mais nous décrivons ses interactions avec syntaxe et "pragmatique"².

1.3 Contraintes

Nous avons choisi de nous fixer des contraintes de **parcimonie** et de **plausibilité cognitive**. Nous insisterons davantage sur la parcimonie pour évaluer nos travaux et les autres modèles concurrents. La plausibilité cognitive reste une notion relativement floue et constituera pour nous plus un cadre général ou une direction, plutôt qu'une contrainte rigoureuse. Ces contraintes ont été sélectionnées dans le prolongement des travaux de Ghadakpour (2003) et du cadre plus général de la théorie algorithmique de l'information (Li & Vitanyi, 1997).

La **parcimonie**, ou minimalisme, repose sur l'évaluation de la complexité d'un objet, comme par exemple un algorithme. Cela signifie que plus la description minimale d'un objet est petite plus l'objet sera dit parcimonieux (en comparaison avec les descriptions d'autres objets). Cette idée de minimalisme est générale. L'une de ses traductions les plus connues est donnée par le rasoir d'Ockham. Celui-ci énonce que "les hypothèses suffisantes les plus simples sont les plus vraisemblables".

La prise en compte de cette contrainte de parcimonie nous a conduit à aborder certains des travaux sur le temps et l'aspect avec un regard plus critique, en ne nous focalisant pas tant sur les résultats des modèles mais plutôt sur la parcimonie de leur(s) algorithme(s). Notre propre modèle est conçu pour rester le plus parcimonieux possible, tout en identifiant les parties qui n'ont pu être simplifiées³.

2 Nous éviterons de faire appelle au terme de pragmatique trop souvent car celui-ci englobe, selon les définitions, un panel très (trop) large de phénomènes linguistiques. Nous en parlons un peu plus dans le chapitre 4.5 Pragmatique, page 65.

3 "simplifié" au sens de "réduire sa complexité de description".

La **plausibilité cognitive** est moins une contrainte qu'une indication du chemin à suivre. Nous avons fait plusieurs hypothèses sur la manière dont le cerveau humain traite l'aspect (et le langage en général) en veillant à ce que notre propre modèle soit compatible avec ces hypothèses : pauvreté du stimulus, pas d'algorithmes attachés au lexique, modularité du système langagier.

Nous sommes conscient que, contrairement à la parcimonie qui est quantifiable, la plausibilité cognitive reste un principe plus flou et les hypothèses faites ne sont pas partagées par tous les chercheurs.

1.4 Modèles existants sur le temps et l'aspect

Il existe déjà un grand nombre de modèles pour le traitement de la temporalité dans les langues naturelles, qui traitent non seulement du temps mais aussi d'aspect et de modalité. Notre objectif n'a pas été de les reprendre pour les confronter ou en réaliser des améliorations, mais surtout d'en tirer des critiques qui nous permettront de contraindre notre propre modèle et de les observer au travers de la lentille de la parcimonie.

Nos critiques et remarques ne porteront pas sur les résultats de ces modèles, dont certains abordent à peine l'aspect, mais sur **comment** les résultats sont obtenus. On ne veut pas comparer ces approches entre elles, mais en tirer des contraintes et des éléments de solution.

De **Ghadakpour**, nous suivrons la volonté de parcimonie et l'idée d'un "module perceptif". On gardera surtout de son ébauche de modèle du traitement de la temporalité la notion de point de vue (viewpoint).

De **Reichenbach**, nous montrerons le caractère dynamique et parcimonieux de son modèle pour le traitement du temps ; Puis supposerons que ce qu'il a fait pour le temps est transposable à l'aspect.

Les contraintes de parcimonie et de plausibilité cognitive que nous nous sommes fixées, en partie grâce à l'étude des deux auteurs précédents, vont nous permettre d'aborder et critiquer des

modèles plus généralistes :

De **Gosselin**, nous critiquerons surtout la possibilité d'avoir des instructions attachées aux lexèmes et grammèmes ; ainsi que les risques d'inflation de la mémoire liée à l'absence d'opération "d'oubli".

À propos de **Kamp** et **Reyle**, nous nous inspirerons de leur structure, la Discourse Representation Structure, et garderons le principe d'unification de structures.

Nous nous intéresserons à la démarche originale et parcimonieuse de **De Swart**, tout en émettant une critique sur la dilution des concepts de telicity et de viewpoint.

1.5 Points clés de notre approche

Les contraintes choisies et certaines des hypothèses formulés nous ont conduit à adopter des approches particulières qui nous distinguent en partie des modèles du traitement du temps et de l'aspect :

- Nous faisons l'hypothèse de **fusions** syntaxique et sémantique simultanées.
- La fusion sémantique a besoin d'un couple ordonné d'éléments, ce qui implique une relation étroite avec une **syntaxe** capable de construire des **arbres binaires**.
- Afin de traiter d'aktionsart, des durées typiques des événements et de la mémorisation de l'ordonnement des situations, nous postulons l'existence d'un **module perceptif** distinct des ontologies classiques (Ghadakpour). Le fonctionnement du module perceptif ne sera pas décrit, mais les interactions avec ce module le seront.
- Les algorithmes nécessaires pour le traitement de l'aspect sont attachés au mécanisme de **fusion** : ce sont les **opérations aspectuelles**. Aucun algorithme n'est attaché au lexique ou à la syntaxe.
- La **parcimonie** du modèle est contrôlée grâce à un nombre fixe d'algorithmes non-récurrents (les opérations aspectuelles), une structure unique de taille fixe : l'**AIS** (Aspectual Information Structure), et un mécanisme d'oubli : la fusion sémantique.

1.6 Choix des énoncés

Le modèle a été élaboré sur un petit nombre d'exemples choisis : Certains des énoncés sont issus des travaux des modèles existants, d'autres ont été imaginés pour mettre en avant l'aspect et limiter au maximum des influences non-aspectuelles. D'autres encore ont été ajoutés pour compléter les possibilités de combinaison des éléments influant sur l'aspect.

La plupart des exemples cités dans ce document peuvent être correctement traités par notre modèle, même si seulement un peu plus d'une cinquantaine d'exemples ont été testés exhaustivement ([14.1 Listes d'exemples et de leurs interprétations](#)).

2 Contraintes

2.1 La parcimonie

Avant de détailler la notion de parcimonie, il nous semble important de préciser ce que nous entendons par le terme de "modèle", ces deux notions étant intimement liées.

2.1.1 Le modèle scientifique

Nous cherchons à proposer un modèle du traitement de l'aspect dans les langues naturelles, mais qu'entendons-nous par "modèle" ? La description qui suit est inspirée de Jean-Louis Dessalles (Dessalles, 2008).

Le mot modèle est polysémique même dans son acception scientifique. Les deux notions les plus courantes sont celles de modèle logique et de modèle scientifique.

- modèle en logique : base de faits (du monde ou d'un monde) et de relations sur ces faits qui permet d'interpréter un ensemble de formules mathématiques. Ainsi, les nombres et leur multiplication, ou les ensembles et leur intersection, sont deux modèles permettant d'interpréter des formules du type $x \times y = z$, $x \times y = y \times x$, $x \times 1 = x$.
- modèle scientifique : un modèle scientifique est une simplification d'un ensemble de faits du monde perçu.

Noter que ces deux notions fonctionnent "à l'envers". La notion de modèle en logique a été introduite pour généraliser la notion de "vérité". Ainsi, un ensemble de formules (ou d'assertions) peut être vérifié dans un modèle et pas dans un autre. Ainsi, pour tout x il existe y tel que la formule $x \times y = 1$ est vraie dans l'ensemble des nombres (non nuls) mais sera fausse en général

dans l'interprétation où \times signifie l'intersection des ensembles.

Types de modèles scientifiques

Le type de modèle le plus élémentaire est une classification (taxonomie). Une classification arborescente interdit à une donnée d'être dans plusieurs classes. Un modèle plus élaboré comportera des contraintes logiques plus sophistiquées ; par exemple des contraintes d'antériorité ou de précédence, ou des contraintes d'exclusion. Les modèles les plus aboutis comportent des lois quantitatives qui relient des grandeurs par des relations valables pour des ensembles continus de valeurs.

Qu'est-ce qu'un bon modèle scientifique ?

La discussion concernant les qualités que doit posséder un bon modèle a progressé au cours du XX^e siècle. On a progressivement compris qu'un bon modèle doit être cohérent, réfutable, prédictible et parcimonieux.

- La notion de cohérence logique recouvre la cohérence interne (absence de contradiction) et la cohérence avec les observations.
- Les notions de réfutabilité et de prédiction sont proches : la notion de réfutabilité, introduite par Popper (1982), nous rappelle qu'une théorie qui n'est pas suffisamment contrainte ne peut certes être mise en défaut, mais n'est aussi d'aucune utilité. Un modèle doit prendre des risques pour être utile.
- Un bon modèle doit par ailleurs faire un maximum de prédictions. On ne distingue pas, en général, les notions de pré-diction et de post-diction (car la distinction n'est pas relative à la théorie mais à l'état épistémique, c'est-à-dire la valeur d'opinion/croyance/certitude, de son juge).
- La notion de parcimonie a conduit à redéfinir la notion de modèle.

La définition moderne de ce qu'est un "bon" modèle subsume toutes les propriétés précédentes.

*Le meilleur modèle est celui qui réalise
la plus grande compression des observations.*

Selon l'aphorisme de Gregory Chaitin (2004), "Comprehension is compression". Le mot compression fait référence à une diminution de complexité (au sens de Kolmogorov). Le meilleur

modèle est ainsi celui qui permet de résumer au mieux un ensemble d'observations (actuelles ou potentielles). Cette définition entraîne :

- la cohérence interne et la réfutabilité (un modèle incohérent ou irréfutable est compatible avec tout et ne comprime rien).
- la cohérence externe et la prédictibilité (si des données restent inexplicées par le modèle, elles ne sont pas comprimées).
- la parcimonie (les hypothèses du modèle s'ajoutent à la compression des observations).

On peut ajouter aux qualités d'un bon modèle la vertu d'être esthétique. Cette qualité coïncide avec celle de simplicité. La définition "compressiste" de la notion de modèle est celle que nous retiendrons par la suite.

2.1.2 La complexité de Kolmogorov

L'une des heuristiques fondamentales en science est le **rasoir d'Ockham**⁴, formulée par le philosophe franciscain Guillaume d'Ockham mais probablement connue avant lui. Le rasoir d'Ockham énonce que "les hypothèses suffisantes les plus simples sont les plus vraisemblables", ou, en d'autres termes, "que les entités ne doivent pas être multipliées au-delà de ce qui est nécessaire". Si le terme de rasoir d'Ockham est le plus populaire, on peut aussi retrouver cet énoncé sous le terme de **principe de simplicité**, **principe d'économie** ou **principe de parcimonie**. C'est ce dernier terme que nous retiendrons ici pour guider notre démarche, afin de soutenir les critiques de modèles du traitement temporel existant, mais aussi pour la conception d'un nouveau modèle.

Voici trois séquences afin d'illustrer ce que signifie la **complexité** dite de Kolmogorov (Li & Vitanyi, 1997) :

- (a) 1100110011001100110011001100110011001100110011001100110011001100
- (b) 0010010000111111011010101000100010000101101000110000
- (c) 0010100101110110101101000101101011011010001011001011

⁴ L'induction de Solomonoff (1964) en est une formalisation mathématique et une preuve.

Quelle est la séquence la plus complexe ? La moins complexe ?

L'exemple (a) est facile à décrire : c'est une suite de treize occurrences de la chaîne '0011'. La description de (b) et (c) est moins facile à déterminer, mais on peut apprendre que (b) est en réalité une représentation binaire des premiers nombres de Pi. La séquence (c) est quant à elle une tentative (manuelle) de répartition "aléatoire" de nombres binaires.

La complexité d'un objet s se définit comme (d'après Li & Vitanyi, 1997) : "la taille du plus petit programme qui donne s ". D'autres définitions moins formelles pourraient être "la taille de la plus petite description de s " ou "la taille du meilleur résumé de s ".

La complexité de description de s est la taille du plus petit programme qui produit s en sortie lorsqu'il est exécuté sur une machine de Turing universelle M :

$$C_M(s) = \min_p \{ |p| : M(p) = s \}$$

Parler de parcimonie ou de minimalisme consiste à déterminer la complexité d'un objet s : un algorithme parcimonieux correspond à un algorithme de faible complexité. Nous voulons **quantifier la parcimonie** d'un modèle du traitement du langage. Nous ne nous Mais avant cela – en plus de l'argument du rasoir d'Ockham – voyons pourquoi cette parcimonie est importante.

Prenons m énoncés, dans une langue quelconque, et grammaticalement corrects.

m énoncés

Peu importe leur forme ou leur longueur pour notre démonstration, mais un énoncé possible pourra par exemple être :

(24) *Un jour, un philosophe sera roi.*

Chacun des énoncés peut avoir zéro ou plus interprétations possibles, sachant que ces interprétations peuvent être plus ou moins précises et dépendantes du contexte. Pour l'exemple (24) la logique temporelle propose plusieurs interprétations (voir page 34 pour les détails) : qu'il existe aujourd'hui un philosophe qui deviendra roi dans le futur, qu'il existera quelqu'un qui deviendra philosophe puis roi plus tard, etc.

On se contentera pour le moment de dire que les m énoncés donnent lieu au total à N interprétations différentes.

N interprétations

Nous voici donc avec nos listes d'énoncés et d'interprétations. Nous allons supposer dans un premier temps que nous connaissons déjà les interprétations de chaque énoncé, et que ces informations se trouvent stockées dans une mémoire. La structure informationnelle la plus petite que nous allons utiliser est le bit, qui est une variable binaire pouvant prendre les valeurs 0 ou 1. Les informations auront cette présentation :

Tableau 1 : Relation entre énoncés et interprétations

Énoncé 1	Interprétation 1	0 (non) ou 1 (oui)

	Interprétation N	
Énoncé 2	Interprétation 1	
	...	
	Interprétation N	
...		
Énoncé M		

C'est le cas où il n'existe pas de modèle (null-model). Les solutions sont connues et aucun calcul n'est présent. Le nombre S de bits nécessaires pour stocker les solutions est égal à :

$$S = M \times \log_2(2^N) = M \times N$$

Le nombre d'énoncés possibles d'une langue étant potentiellement infini, ainsi que leurs interprétations possibles, le nombre de bits nécessaires pour interpréter un énoncé serait lui aussi infini ! Voici pour un cas extrême. Dans le cas où chaque exemple ne peut avoir qu'une seule interprétation, il faut $\log_2(N)$ bits pour discriminer chaque interprétation (vrai ou faux). Soit un nombre de bits pour M énoncés égal à :

$$S = M \times \log_2(N)$$

Dans l'idéal nous cherchons un modèle de taille constante, c'est-à-dire qui est capable de gérer un nombre infini d'énoncés (et est donc indépendant de M) tout en fournissant toutes les interprétations possibles (et est donc indépendant de N).

En pratique, le "modèle parfait" pour le traitement temporel dans le langage n'existe pas (on ne le

connaît pas en tout cas), tout comme le null-model d'ailleurs (nous ne disposons pas d'une infinité de bits, ni des moyens pour créer une telle base de données). Mais ces deux situations nous ont permis de montrer les cas les plus extrêmes du point de vue de la parcimonie : le "modèle parfait" est le plus parcimonieux possible puisqu'il ne demande aucune information supplémentaire par rapport aux énoncés, tandis que le null-model n'est absolument pas parcimonieux.

En prenant en compte ce principe de parcimonie, les modèles de traitement du langage doivent être capables de quantifier le nombre de bits nécessaires à leur fonctionnement. Les grammairiens ont classé les lemmes en différentes classes : nom, verbes, pronoms, etc. et mis au point des règles de grammaire. Ces grammaires se sont enrichies (tout en devenant moins parcimonieuses) pour pouvoir traiter davantage de cas.

Prenons par exemple la classification des trois groupes de verbes en français. Chaque verbe appartient à l'un des trois groupes, les verbes du troisième groupe disposant à leur tour de sous-groupes et d'exceptions (les exceptions coûtent cher en termes de parcimonie puisque pour Q verbes irréguliers on a besoin de $Q \times \log(N)$ bits supplémentaires) ! De plus, chaque forme irrégulière de verbe fait à son tour augmenter le nombre d'interprétations N .

Les noms en français peuvent eux aussi être catégorisés : par leur genre (féminin ou masculin), leur nombre (singulier ou pluriel), etc.

Attardons-nous sur le genre des noms en français. C'est une information qui ne nécessite qu'un seul bit (0 pour le féminin et 1 pour le masculin, par exemple)⁵. Si le modèle connaît le genre de tous les noms, et que l'on considère qu'il existe K noms connus dans la langue, alors on dira que le modèle utilise K bits. K peut être très grand, mais sera néanmoins fini (à un instant donné). Ainsi, si M (le nombre d'énoncés possibles) tend vers l'infini, K peut être limité, et le modèle devient plus parcimonieux.

On se rend compte que chaque "trait", tel que le genre, fait augmenter sensiblement le nombre de bits nécessaires au modèle. Dans le cas des noms ou des verbes, catégories très vastes (plusieurs milliers de membres) mais aux valeurs binaires⁶, chaque nouveau trait assigné multiplie par K le

5 Il se trouve que le français a des exceptions qui peuvent prendre les deux genres, comme *un bel orgue/de belles orgues*. Par soucis de simplification, nous ignorerons ces cas.

6 L'hypothèse de traits binaires est ici faite pour simplifier la démonstration,.

nombre de bits nécessaire (K étant ici le nombre de noms connus).

Il est bien sur envisageable d'avoir une compression de l'information et une réduction des bits nécessaires, mais le risque d'inflation est toujours présent et doit toujours être pris en compte.

Il existe des lemmes dans les langues naturelles qui appartiennent à ce que l'on appelle les "classes fermées" : ces lemmes sont en nombre limité et sont généralement stables, c'est-à-dire que de nouveaux éléments ne sont pas créés régulièrement (comme c'est le cas pour les noms et verbes). Victorri distingue les mots du lexique (noms, verbes, adjectifs) qu'il place dans la "classe ouvertes" et les mots de la grammaire (prépositions, conjugaisons, etc.) qu'il place dans l'une des "classes fermées". Par exemple, *pendant* est une préposition qui appartient à cette "classe fermée" sur le temps. C'est aussi le cas, par exemple, de *depuis*, *hier*, ou de la conjugaison *-ait*. Ces lemmes sont connus et peuvent être listés. Ce sont eux qui introduisent la plus grande partie de l'information temporelle d'un énoncé.

Parce qu'ils sont peu nombreux (moins d'une centaine pour le français, en considérant chaque temps grammatical comme un lemme), le nombre de bits utiles que les lemmes de la classe fermée utiliseront le sera (a priori) aussi.

On peut choisir par exemple d'assigner à un élément du lexique un trait "relation" qui pourra prendre les valeurs "simultané" ou "précède". *Pendant* aura la valeur "simultané", tandis que *avant* aura la valeur "précède".

En traitant certains exemples, on se rendra compte que ce trait de relation n'est pas suffisant, pour différencier la simultanéité parfaite (lorsque les débuts et fins de deux intervalles coïncident) et l'inclusion. Par exemple :

(25) *Elle a parlé pendant dix minutes* (exactement).

(26) *Elle a parlé pendant le spectacle* (exactement OU à un moment précis OU à plusieurs reprises).

Le trait binaire des "relations" s'avère ici insuffisant puisqu'on voudrait exprimer l'inclusion ou la simultanéité parfaite. Comment résoudre ce problème ? Faut-il ajouter des valeurs aux traits et perdre la binarité ? Ajouter un nouveau trait ? Ne pas ajouter de trait à *pendant* mais obtenir l'information à partir d'un calcul global prenant en compte *pendant* et le circonstanciel (*dix*

minutes et le spectacle) ? Voir assigner un algorithme directement à pendant ?

Le choix n'est pas facile, surtout que chacun peut nous permettre d'atteindre notre objectif. Le choix le plus facile serait de fournir un maximum d'information afin de traiter tous les cas possibles. Seulement... nous perdons en parcimonie. Mais en cherchant à rester trop parcimonieux, nous prenons le risque de ne plus traiter tous les cas.

Nous avons parlé jusqu'à présent de "traits" assignés aux lemmes, c'est-à-dire de caractéristiques des éléments du lexique : ce seront les informations en entrée.

Nous choisissons de ne pas prendre en compte les règles grammaticales dans ce calcul puisqu'elles seront traitées par un module indépendant qui ne fait pas partie du modèle du temps. On comptera néanmoins comme informations en entrée (input) les données qui proviennent du module perceptif : Ordre de grandeur typique et télélicité notamment (voir chapitre [6.3 Le module perceptif](#), page 94).

2.2 La plausibilité cognitive

Nous avons parlé dans le chapitre précédent de parcimonie, en invoquant le rasoir d'Ockham pour justifier son utilisation. Il existe une seconde contrainte que nous avons voulu garder à l'esprit pendant la thèse : la plausibilité cognitive.

Ce que nous appelons plausibilité cognitive est l'idée de rester compatible avec le fonctionnement supposé (du cerveau) d'un être humain. Puisque le sujet d'étude est le langage naturel et son interprétation, c'est le comportement humain, ou en tout cas la représentation que nous nous en faisons, qui est l'exemple à suivre.

Cette contrainte de plausibilité cognitive est malheureusement vague et difficile, voire impossible, pour le moment à quantifier. Rien ne nous permet d'affirmer avec certitude que tel modèle est plus "cognitif" qu'un autre. Néanmoins, divers éléments peuvent être mentionnés pour soutenir notre démarche.

Commençons cette réflexion sur les caractéristiques d'un module du traitement temporel par une observation de tous les jours. C'est dans la vie quotidienne qu'enfants et adultes font l'expérience de la communication langagière. Les gens apprennent la même chose dans un même environnement, y compris les "fautes" par rapport à une éventuelle langue de référence. Cet apprentissage converge et donne lieu à peu ou pas de variations. Cette observation nous indique que le modèle de traitement du langage, malgré la puissance expressive que nous lui prêtons dans les paragraphes précédents, doit aussi être un modèle **contraignant**. Il ne doit pas exprimer plus de choses que ce que l'on peut observer. Ainsi tout modèle qui cherche à traiter du langage doit rejeter les énoncés que les locuteurs considèrent comme "bizarres" ou "incorrects".

Un argument cognitif en faveur de la parcimonie du système langagier est une conséquence de ce que Chomsky a appelé la "pauvreté du stimulus" (Chomsky, 1988; R. Jackendoff, 2002). Cet argument consiste à dire que l'enfant n'est pas exposé à une quantité suffisante d'énoncés pour mettre en place un modèle statistique (Crain, 1991), comme cela se fait actuellement dans de nombreux modèles de traitement du langage.

L'argument ne saurait à lui seul imposer la parcimonie car il n'empêche pas l'existence d'un appareil langagier inné qui, lui, ne serait pas parcimonieux. Que la complexité soit innée acquise, elle semble cependant incompatible avec la nécessité pour les enfants de développer une compétence dans une langue donnée. Imaginons un mécanisme complexe, par exemple un système dans lequel des algorithmes sont attachés au lexique⁷. Chaque mot (potentiellement) pourrait ainsi disposer d'une suite complète d'instructions. Il est impossible que les enfants possèdent tous ces algorithmes dès la naissance : Ils sont capables d'apprendre n'importe quelle langue, et pas seulement celle de leur parents.

Cela veut dire que les enfants, ou les apprenants en général, devraient acquérir ces algorithmes en observant leur environnement linguistique ; et cela pour chaque élément lexical ! Or la lenteur de l'apprentissage et les risques d'erreurs augmentent considérablement avec la complexité de ce qu'il est nécessaire d'acquérir. Pourtant, les locuteurs d'une même langue, dès leur plus jeune âge, se comprennent sans trop de problèmes.

On ne prouvera pas ici qu'il est impossible d'associer des algorithmes à des éléments du lexique, mais nous disons qu'un modèle qui réussit à s'en passer sera beaucoup plus simple, et donc

⁷ C'est-à-dire qui puisse exécuter des instructions précises lors de la reconnaissance de lemmes particuliers.

plausible, qu'un modèle qui imposerait aux apprenants d'acquérir une structure algorithmique nouvelle pour chaque signification apprise.

Revenons sur la notion de parcimonie pour un appareil langagier. Une solution de type "inné" serait de supposer qu'il existe une "sémantique universelle", sorte de méta-sémantique, ou langage de la pensée (Fodor, 1983), au même titre que la "grammaire universelle" qui, selon certaines théories, supporterait toutes les grammaires humaines. Une telle sémantique universelle peut-elle exister et quelle forme aurait-elle ?

La recherche d'un déterminant universel pourrait être comparable en syntaxe (Chomsky, 1995) et en sémantique (R. Jackendoff, 2002), mais dans cette hypothèse on ne ferait que déplacer le problème : Quel serait le modèle le plus minimaliste de ce méta-langage, s'il existe ?

Pour essayer de d'estimer la complexité de cette sémantique universelle (traitement universel de la sémantique), il faut, à notre avis, chercher dans d'autres domaines que celui de la linguistique. Prenons le cas de l'évolution. Si l'on compare aujourd'hui les différences génétiques d'un être humain avec l'un de ses plus proche cousin, le chimpanzé qui, lui, n'est pas doté des mêmes capacités langagières que les humains, on s'aperçoit que ces différences sont très faibles (Mikkelsen et al., 2005) ; suffisamment faibles pour douter qu'il soit possible d'y encoder un appareil linguistique fonctionnel complexe. Si l'on considère comme tout à fait possible pour cet appareil de réutiliser des parties plus anciennes du cerveau, on considère que la complexité de ce qu'on est obligé de postuler ne peut excéder la complexité de la différence génomique. La démarche inverse reviendrait à supposer un appareil langagier entièrement fonctionnel existant chez le chimpanzé, mais non exprimé, ce que l'on se refuse à imaginer.

De plus, le langage n'est pas la seule différence existant entre humain et chimpanzé : il semblerait que nos cousins primates aient une capacité à se projeter dans l'avenir différente de celle de l'humain (Mulcahy & Call, 2006; Roberts, 2007).

Que pouvons-nous déduire de ces faibles différences génétiques entre humains et chimpanzés ? Que la capacité langagière, qui est l'un des éléments qui nous différencient, est certainement le résultat d'un petit nombre de gènes, nombre largement insuffisant pour encoder une compétence langagière qui serait complexe. Pourtant cette différence est suffisante pour que les humains puissent développer leur capacité langagière au cours de leur vie, ce qui inclut le traitement du temps, de l'espace, de la causalité, de déduction, etc.

Comment cela a-t-il pu se passer ? Rappelons que c'est l'évolution qui a "conduit" à ce résultat. Les humains peuvent parler, mais notre ancêtre commun avec le chimpanzé ne le pouvait pas. Pourtant, on peut penser que celui-ci était bien adapté à son environnement et que son cerveau devait être parfaitement fonctionnel. Le cerveau d'*Homo Sapiens* ne s'est certainement pas reconfiguré entièrement et dispose toujours des anciennes parties. Ce sont les nouvelles parties, celles qui sont apparues depuis, qui nous intéressent. On suppose que ces parties - appelons-les "modules" pour faire une analogie avec l'informatique (Fodor, 1983) - ne sont pas très complexes (sinon elles ne seraient pas apparues au cours de l'évolution, ni ne tiendraient dans les quelques gènes qui en seraient responsables) et qu'elles n'ont pas beaucoup de connexions avec les autres modules (pour les mêmes raisons). L'apparition du langage chez l'être humain est certainement le fruit de mutations, donnant naissance à de nouveaux modules (ou à des modifications de modules déjà existants), voire à un seul module.

Pour reformuler, nous supposons que pour être compatibles avec l'évolution, le ou les modules traitant du langage sont peu nombreux et "simples". Notre idée est de considérer comme synonymes de ce "simple" notre définition de la parcimonie. La plupart de ces modules existaient déjà avant l'apparition du langage chez l'humain, et le langage n'est peut-être que le résultat d'un nouveau module, ou de la connexion de modules existants (Miyagawa, Berwick, & Okanoya, 2013).

Nous avons parlé de système langagier, mais ce qui nous intéresse plus précisément est bien évidemment le système d'expression et d'interprétation **temporelle**⁸ dans le langage. Celui-ci est peut-être une partie d'un module de gestion de la topologie ou d'un module pour se projeter dans le futur (Núñez & Cooperrider, 2013; Suddendorf & Corballis, 2007). Nous ne démontrerons pas ici que le(s) module(s) du traitement du temps (temporalité, aspect, modalité) est (sont) unique(s) à l'être humain, mais nous garderons l'idée que ce ou ces modules doivent certainement être parcimonieux.

Un autre argument sur la plausibilité cognitive est celui de la granularité : Ghadakpour (2003) a traité la question de la structure temporelle dans la mémoire humaine et considère que la mémoire temporelle ne peut pas disposer d'un ordre total, c'est-à-dire que n'importe quelle situation A ne pourra pas être ordonnée par rapport à une autre situation B. Que les "bornes" de

8 Système qui inclurait le traitement de l'**aspect**.

début et de fin soient antérieures, postérieures, ou simultanées aux "bornes" de B, il n'est pas possible de pouvoir **comparer** n'importe quelles situations entre elles, y compris dans le cas de situations conditionnelles.

D'après Ghadakpour, si la structure temporelle existe celle-ci doit être dense, c'est-à-dire non granulaire⁹, et donc infinie. Ce qui est absurde.

L'exemple typique pour illustrer ce problème est celui dit de Zénon. Zénon, un philosophe de l'antiquité grecque, a proposé un exercice de pensée pour cela : imaginons un archer tirant une flèche sur une cible. Il y a un moment où la flèche atteindra la moitié de sa distance. On peut alors s'imaginer que la flèche va parcourir la moitié de la distance restante. Puis de nouveau, imaginons-nous que la flèche parcourt à nouveau la moitié de la distance restante (un quart de la distance totale donc), et de nouveau. Et ainsi de suite. Il sera toujours possible pour un être humain de s'imaginer un nouveau moment inséré entre l'impact et le moment imaginé précédemment. Heureusement pour eux, les gens se fatiguent vite de cet exercice de pensée, mais un algorithme fonctionnant de la même manière et ne connaissant pas la fatigue se retrouverait bloqué dans une boucle infinie...

Le cerveau humain ne contiendra évidemment jamais assez de neurones pour stocker tous les moments imaginables : La représentation physique (dans un cerveau, ou dans un ordinateur) de cette infinité de moments est impossible.

Un autre exemple pour soutenir l'absence de granularité de la structure physique de la mémoire temporelle est donné par Ghadakpour :

(27) *Il y a quinze milliards d'années, trois secondes après le début de l'univers, la symétrie entre la matière et l'antimatière s'est brisée.*

Cette phrase n'est pas trop difficile à comprendre. Même en ignorant la signification de *la symétrie entre la matière et l'antimatière* (tout le monde n'est pas physicien), il nous est possible d'ordonner cette situation avec *le début de l'univers*.

Si on suppose que les représentations des situations sont positionnées sur une ligne temporelle, avec une certaine échelle et une certaine précision, la mémoire temporelle devrait disposer d'au moins un demi-milliard de milliards d'éléments représentant chaque seconde. Nous sommes déjà à trois ordres de grandeurs au-dessus du nombre de synapse d'un cerveau (Ghadakpour, 2003).

⁹ Au sens de "grain minimum élémentaire", "d'atomes".

D'après Ghadakpour, il n'existe qu'une seule solution pour représenter une structure qui respecte le raisonnement de Zénon sur une mémoire physique : elle consiste à réutiliser des emplacements mémoires. Cela signifie qu'une partie de nos représentations sont temporaires et que le stockage des informations sur les situations n'est pas systématique.

Nous suivrons dans ce document les conclusions de Ghadakpour, à savoir qu'il n'existe pas de représentation granulaire du temps dans la mémoire humaine, mais qu'il est impératif de disposer d'un système d'ordre relationnel entre la représentation des situations.

Nous avons présenté dans ce chapitre ce que nous entendons par "plausibilité cognitive" : c'est-à-dire la prise en compte des connaissances sur l'évolution humaine et de l'apprentissage du langage pour se fixer des contraintes dans la conception d'un modèle.

Nous ne souhaitons pas réaliser de système à base de connaissances ou de programme statistique complexe pour le traitement du temps dans la langue (qui pourraient être toutefois très efficaces et amplement suffisants pour de nombreuses tâches), mais proposer une ébauche d'un modèle parcimonieux et cognitivement plausible de cette tâche chez l'humain.

3 Approches non linguistiques du temps

Le temps est un "objet" étrange. Objet ou concept, les humains s'y réfèrent en permanence dans la vie de tous les jours. Puisque "le temps passe", les choses changent. Mais est-ce que le temps existe vraiment, où n'est-ce qu'un concept abstrait développé par l'humain pour rendre compte des changements sur lui-même et dans le monde ?

La question a déjà été posée par les physiciens. Ceux-ci utilisent la variable temporelle T , sans jamais avoir pu identifier ou isoler ce que le temps est exactement, autrement que par l'observation de changements dans le monde (voir E. Klein, 2004). Nous ne parlerons pas du temps lui-même dans ce chapitre, mais du **traitement du temps** chez les humains, et plus spécifiquement de ses manifestations dans les langues naturelles. Au-delà des approches en philosophie, en logique, et en linguistique cognitive, que nous aborderons de manière superficielle, nous présenterons un catalogue des phénomènes temporels étudiés par la **linguistique**.

3.1 En philosophie

Avant les physiciens modernes, les philosophes de toutes les époques se sont attaqués au concept du temps, sans jamais arriver à une conclusion satisfaisante. Pour reprendre les mots célèbres de Saint Augustin (Augustin, 1993) :

Qu'est-ce que en effet que le temps ? Qui saurait en donner avec aisance et brièveté une explication ? ... Si personne ne me pose la question, je le sais ; si quelqu'un pose la question et que je veuille expliquer, je ne sais plus.

Aujourd'hui, il existe encore deux écoles philosophiques à propos de la nature du temps : les

réalistes qui pensent que le temps a une existence propre, indépendante de l'esprit humain ("héritiers" de Newton et de ses variables temporelles, puis plus directement de S. Clarke), et les idéalistes qui pensent le contraire ("héritiers" de G. W. Leibniz)¹⁰.

C'est une simplification bien sûr, et nous éviterons de rentrer dans le détail. Le débat n'est néanmoins pas anodin et il nous semble utile de le mentionner pour pouvoir répondre au problème de la représentation du temps dans la mémoire humaine. Précisons que nous parlerons alors du concept de temps, sa représentation cognitive, c'est-à-dire du temps tel qu'il est représenté dans le cerveau humain, et pas de son existence comme dépendant ou non de l'esprit humain.

La question est donc de connaître la structure de représentation du temps des êtres humains. Comment savons-nous qu'Alexandre le Grand précède Jules César, ou que la première guerre du Golfe s'est déroulée pendant la présidence de François Mitterrand ? Un système d'ordonnement semble une condition nécessaire, mais le temps lui-même est-il "matériellement" représenté ? Autrement dit, la ligne temporelle que nous traçons sur une feuille de papier lorsque l'on souhaite schématiser le temps a-t-elle une existence dans notre cerveau ?

3.2 En logique

Ce chapitre est en partie inspiré des travaux d'Antony Galton¹¹.

La logique temporelle englobe les approches de la représentation des informations temporelles dans le cadre de la logique formelle, et plus spécifiquement de la logique modale (Prior, 1967, 1968). Grâce à sa puissance et son formalisme, la logique temporelle est utilisée dans des domaines aussi variés que la philosophie, la linguistique (sémantique), l'intelligence artificielle et l'exécution des logiciels informatiques. La logique temporelle est-elle adaptée à notre sujet ? Et si non pourquoi ?

Voyons d'abord en quoi elle consiste.

¹⁰ Voir Leibniz & Clarke, 2000 pour la correspondance entre les deux philosophes ; ou <http://www.earlymoderntexts.com/leibclar.html> pour une version téléchargeable.

¹¹ <http://plato.stanford.edu/entries/logic-temporal/>

La logique temporelle, ou Tense Logic, dispose de quatre opérateurs modaux supplémentaires par rapport à la logique modale (il est possible de se restreindre à deux opérateurs) :

P "Il a été vrai à un moment que ..."

F "Il sera vrai à un moment que ..."

H "Il a toujours été vrai que ..."

G "Il sera toujours vrai que ..."

Comme exemple d'utilisation, Prior (1967) utilise ces opérateurs pour représenter certaines expressions philosophiques à propos du temps :

$Gp \rightarrow Fp$ "What will always be, will be"

$G(p \rightarrow q) \rightarrow (Gp \rightarrow Gq)$ "If p will always imply q, then if p will always be the case, so will q"

$Fp \rightarrow FFp$ "If it will be the case that p, it will be — in between — that it will be"

$\neg Fp \rightarrow F\neg Fp$ "If it will never be that p then it will be that it will never be that p"

La Tense Logic est obtenue en ajoutant les opérateurs temporels à une logique existante et ses axiomes. Il est possible d'obtenir d'autres systèmes de logique temporelle avec différentes bases logiques. Il est ainsi possible d'exprimer des distinctions d'importance dans l'interprétation d'énoncés en langue naturelle. Par exemple, l'énoncé *A philosopher will be a king* peut être interprété en logique de nombreuses manières :

(28) $\exists x(\text{Philosopher}(x) \ \& \ F \text{ King}(x))$ *Quelqu'un qui est actuellement philosophe sera roi à un moment dans le futur.*

(29) $\exists x F(\text{Philosopher}(x) \ \& \ \text{King}(x))$ *Il existe actuellement quelqu'un qui sera à un moment dans le futur et roi et philosophe.*

(30) $F \exists x(\text{Philosopher}(x) \ \& \ F \text{ King}(x))$ *Il existera quelqu'un qui sera philosophe, et plus tard deviendra un roi.*

(31) $F \exists x(\text{Philosopher}(x) \ \& \ \text{King}(x))$ *Il existera quelqu'un qui sera en même temps philosophe et roi.*

Suivant Prior, la Tense Logic a été régulièrement enrichie pour décrire davantage de situations. Kamp (1968) introduit les opérateurs temporels **S** et **U** ("since" et "until"), avec pour sens :

Spq "q has been true since a time when p was true"

Upq "q will be true until a time when p is true"

Prior introduit en complément des opérateurs la **metric tense logic** : une notation permettant d'exprimer un moment passé ou futur avec un intervalle. Par exemple la notation **Fnp** signifie "It will be the case the interval n hence that p". Pnp n'étant pas nécessaire puisqu'il est possible d'avoir n négatif : F(-n)p. Le présent est exprimé par n=0.

Les opérateurs généraux non-métriques sont alors définissables de cette manière :

$$Pp \equiv \exists n(n < 0 \ \& \ Fnp)$$

$$Fp \equiv \exists n(n > 0 \ \& \ Fnp)$$

$$Hp \equiv \forall n(n < 0 \rightarrow Fnp)$$

$$Gp \equiv \forall n(n > 0 \rightarrow Fnp)$$

Un autre exemple d'ajout est l'opérateur "next time" **O**. Cet opérateur suppose que le temps est composé d'une série de "temps" atomiques (des instants, c'est la vision philosophique des réalistes). La formule Op signifiant alors que p est vrai pour le pas (l'instant) suivant.

Notons que cet opérateur se place dans la perspective d'un temps discret et n'est alors pas adapté au traitement du langage naturel (voir l'exemple de Zénon dans le chapitre Erreur : source de la référence non trouvée Erreur : source de la référence non trouvée, page Erreur : source de la référence non trouvée). Il est par contre très utile dans les logiques de programmation ou l'existence d'un "pas" est avérée (fourni par l'horloge des processeurs). Nous reviendrons un peu plus bas sur ce problème.

La sémantique de la Tense Logic est très proche de la logique modale. Une frame temporelle (temporal frame) consiste en un groupe T d'entités appelées temps et une relation d'ordre "<" dans T. Cela correspond au "flot du temps" dans lesquels le sens des opérateurs temporels doit être défini. Une interprétation du langage de la logique temporelle assigne une valeur de vérité à chaque formule atomique à chaque temps de la frame temporelle. Suivant cette interprétation, on peut donner les définitions suivantes aux opérateurs de tenses :

$$Pp \text{ is true at } t \quad \text{if and only if } p \text{ is true at some time } t' \text{ such that } t' > t$$

$$Fp \text{ is true at } t \quad \text{if and only if } p \text{ is true at some time } t' \text{ such that } t > t'$$

$$Hp \text{ is true at } t \quad \text{if and only if } p \text{ is true at all times } t' \text{ such that } t' > t$$

$$Gp \text{ is true at } t \quad \text{if and only if } p \text{ is true at all times } t' \text{ such that } t > t'$$

De nombreux axiomes ont été suggérés pour exprimer telles ou telles propriétés du flot du temps,

et la sémantique nous donne un moyen précis de définir la correspondance entre des formules logiques temporelles et les propriétés de frames temporelles.

Une formule p est dite "caractériser" un groupe de frames F si :

p est vrai à tout moment sous toutes les interprétations pour chaque frames de F .

Pour toutes frames hors de F , il existe une interprétation qui rend p faux à un moment.

Une formule du premier ordre en $<$ détermine une classe de frames, c'est-à-dire ceux dans laquelle la formule est vrai.

Une formule de logique temporelle p correspond à une formule de premier ordre q tant que p caractérise la classe de frames pour laquelle q est vrai. Voici quelques exemples célèbres de paires de formules :

$H_p \rightarrow P_p$	$\forall t \exists t' (t' < t)$	(unbounded in the past)
$G_p \rightarrow F_p$	$\forall t \exists t' (t < t')$	(unbounded in the future)
$F_p \rightarrow FF_p$	$\forall t, t' (t < t' \rightarrow \exists t'' (t < t'' < t'))$	(dense ordering)
$FF_p \rightarrow F_p$	$\forall t, t' (\exists t'' (t < t'' < t') \rightarrow t < t')$	(transitive ordering)
$FP_p \rightarrow P_p \vee P_p \vee F_p$	$\forall t, t', t'' ((t < t'' \& t' < t'') \rightarrow$ $(t < t' \vee t = t' \vee t' < t))$	(linear in the past)
$PF_p \rightarrow P_p \vee P_p \vee F_p$	$\forall t, t', t'' ((t'' < t \& t'' < t') \rightarrow$ $(t < t' \vee t = t' \vee t' < t))$	(linear in the future)

Il existe néanmoins des formules de logique temporelle (par exemple $GF_p \rightarrow FG_p$) qui ne correspondent à aucune des propriétés de frames de la logique de premier ordre. Et inversement, des propriétés de frames temporelles en logique de premier ordre (comme l'irréflexivité) qui ne correspondent à aucune formule de la logique temporelle (Van Benthem, 1983).

En ce qui concerne la logique des prédicats, la dimension temporelle peut être introduite par l'ajout d'un argument à chaque proposition ou prédicat temporel. Par exemple :

Kill(Brutus, Caesar, -44).

Revenons sur la réification du temps (mentionné avec l'opérateur O). Celle-ci pose problème, philosophiquement mais aussi linguistiquement (En tout cas, si l'on prend en considération l'approche "idéaliste" ou cognitive). Prior (1968) propose de contourner le problème en considérant un instant de cette manière :

‘the conjunction of all those propositions which would ordinarily be said to be true at

that instant'

Les instants sont alors remplacés par des propositions uniques. Une affirmation de la forme "True(p, t)", qui dit que la proposition p est vraie à l'instant t, peut alors être paraphrasée ainsi : " $\Box (t \rightarrow p)$ ", c'est-à-dire l'instant t implique nécessairement p.

Ces manœuvres sont le cœur des **logiques temporelles hybrides** dans lesquelles l'appariement standard des propositions et opérateurs temporels est remplacé par des propositions vraies à des instants uniques, tout en évitant de réifier les-dits instants.

Voici grossièrement pour la logique temporelle, ou tense logic, mais comment la logique traite-t-elle de l'aspect ?

Certains énoncés rapportent des états **homogènes** (voir Chapitre [4.1.8 Homogénéité](#), page 52). Un énoncé tel que *Marie est endormie* est dit homogène, puisque si l'on considère l'énoncé vrai pour chaque intervalle il sera vrai à chaque instant de l'intervalle.

Mais si l'on considère un énoncé décrivant une situation de manière non-homogène, tel que *Jean marche jusqu'à la gare*, celui ne sera plus vrai pour un instant de l'intervalle, ou l'un de ses sous-intervalles. Si *Jean marche jusqu'à la gare* entre midi et midi quinze, l'énoncé est incorrect pour le sous-intervalle midi, midi cinq : Jean sera en train de marcher vers la gare, mais ne l'aura pas atteinte à la fin de l'intervalle considéré.

Ce genre de problème a été résolu en réifiant certaines caractéristiques de l'aspectualité et popularisé en intelligence artificielle (Allen, 1984). Selon cette approche les états (states ; homogènes) et événements (events ; non-homogènes) sont notés comme des termes dans la théorie du premier ordre. Leur incidence temporelle est notée par les prédicats relationnels "Holds" et "Occurs". Par exemple :

Holds(Asleep(Mary), (1pm, 6pm))

Occurs(Walk-to(John, Station), (1pm, 1.15pm))

Avec les deux derniers termes comme les bornes des intervalles.

L'homogénéité (les états), et la non-homogénéité (les événements) sont alors exprimés par les axiomes suivants :

$\forall s, i, i' (\text{Holds}(s, i) \ \& \ \text{In}(i', i) \rightarrow \text{Holds}(s, i'))$

$\forall e, i, i' (\text{Occurs}(e, i) \ \& \ \text{In}(i', i) \rightarrow \neg \text{Occurs}(e, i'))$

"In" exprimant la relation de sous-intervalle.

Un autre problème, dit de "variable polyadicity" apparaît lorsque l'on traite des énoncés disposant d'informations circonstancielles supplémentaires. Par exemple :

(32) *John saw Mary in London on Tuesday.*

(33) *Therefore, John saw Mary on Tuesday.*

Comment représenter en logique des prédicats ces deux exemples ? Davidson (1967) propose de donner à chaque prédicat un argument supplémentaire pour les lier entre eux. La forme logique de (32) et (33) devient alors :

$\exists e(\text{See}(\text{John}, \text{Mary}, e) \ \& \ \text{Place}(e, \text{London}) \ \& \ \text{Time}(e, \text{Tuesday})),$

$\text{Therefore, } \exists e(\text{See}(\text{John}, \text{Mary}, e) \ \& \ \text{Time}(e, \text{Tuesday})).$

Ici, c'est la variable e qui joue le rôle de liant entre les prédicats. C'est cette approche qui est utilisé par exemple dans le modèle computationnel de Kowalski et Sergot (1986).

Nous avons rapidement vu la temporalité, l'aspect, mais qu'en est-il de la modalité ? La logique temporelle la traite en utilisant la ligne du temps et en la "divisant" en certains instants pour représenter les futurs possibles de chacun de ces instants.



Figure 1 : Futurs possibles

La succession des instants qui précèdent l'instant de "l'embranchement" représente alors l'histoire commune des multiples histoires **possibles** (on parle de futurs possibles). Un modèle déterministe de cette représentation est appelé branching future model.

Dans ce type de modèle, les formules sont évaluées en fonction d'un instant et d'une histoire, au lieu de juste un instant. Ainsi, en considérant la paire (h,t) , on interprétera la formule Fp comme vraie tant que p est vraie à un moment dans le futur de l'instant t et déterminé par l'histoire h .

Un nouvel opérateur a été introduit pour tenir compte des futurs possibles : \diamond .

On dira que " $\diamond p$ " est vrai pour (h, t) tant qu'il existe pour une histoire h' tel que " p " est vrai dans (h', t) . " $\diamond Fp$ " signifiera alors que p aura lieu dans un futur possible. L'opérateur " \square " est la version forte de \diamond . " $\square Fp$ " signifiera que " p " est inévitable, c'est-à-dire qu'il aura lieu dans tous les futurs possibles.

Cette interprétation du branching future model est appelé "Ockhamist" par Prior. Une autre interprétation, dite "Peircean", Considère " Fp " comme l'équivalent de " $\square Fp$ ".

3.2.1 Application au langage naturel

Nous avons rapidement passé en revue ce qu'était la logique temporelle, et déjà mentionné qu'elle a des applications dans les domaines de l'intelligence artificielle et de l'informatique. Peut-elle être appliquée au langage naturel ?

Il se trouve que la logique a été utilisée bien avant pour le traitement temporel dans le langage, avec Reichenbach (1947). Nous présentons son modèle plus en détail en page 67. Bien que "logique", son modèle ne ressemble que très peu aux formalismes de la logique temporelle exposés jusqu'à présent : en effet, Reichenbach propose un modèle computationnel, et non pas descriptif. Et justement, la logique temporelle inspirée de Prior a surtout une utilisation descriptive dans le domaine du langage naturel. Par sa puissance, elle permet d'exprimer sans ambiguïtés les sens des énoncés (voir les exemples (28) à (31)) sans devoir passer par du paraphrasage.

Seulement cette puissance est aussi, à notre avis, un défaut quand il s'agit de traiter le langage naturel, et tout particulièrement quand on cherche à être minimaliste. Le langage **est** ambigu et doit le rester ! Il est fort probable qu'aucun lecteur de la phrase *A philosopher will be a king* ne disposera des quatre interprétations que l'on a proposées sans un sérieux effort cognitif. Et pourtant le lecteur affirmera sans aucun doute avoir bien compris la phrase.

(une partie de) La logique temporelle pose un autre problème pour le traitement du langage humain puisqu'elle demande la réification des instants, c'est-à-dire l'existence "matérielle" d'une ligne ou d'un arbre (représentation des futurs possibles) du temps. Or nous allons voir avec Ghadakpour (2003) et le paradoxe de Zénon (page Erreur : source de la référence non trouvée) que

le cerveau humain ne peut pas fonctionner ainsi.

Certains phénomènes spécifiques au langage naturel n'ont d'ailleurs pas ou peu été traités par la logique temporelle : la prise en compte des ordres de grandeurs (page 44) ou le mécanisme de répétition (page 53), par exemple.

3.3 En linguistique cognitive

Un bref historique

La linguistique cognitive s'intéresse au lien entre langage et cerveau sans considérer que le langage est traité par un module indépendant (tel que le fait la grammaire générative (Chomsky, 1995)). Elle s'intéresse aux relations entre la structure du langage et les choses qui en sont indépendantes, tel que les principes de catégorisation, de pragmatique, et les principes fonctionnels comme l'économie ou l'iconicité.

Les approches cognitives du langage sont extrêmement variées, mais elles s'entendent en général pour reconnaître que la chose la plus importante pour comprendre le langage est d'étudier le **sens** qu'il véhicule, et ne pas le traiter (tel que l'a fait à l'époque l'approche chomskienne) comme un résultat secondaire de l'étude du langage dont le pivot serait la syntaxe.

La linguistique fonctionnelle (Bybee, Comrie, Haiman, Hopper, Thompson, Givon) apparaît à peu près à la même époque en se concentrant sur les principes dérivés du langage qui est essentiellement traité comme un moyen de communication. Le langage est alors étudié surtout en situation de discours. Tout comme la linguistique cognitive, elle reconnaît que le meilleur moyen d'étudier le langage est de le placer en contexte, qu'il soit cognitif ou social, et qu'il ne faut pas se limiter au système linguistique proprement dit (c'est-à-dire lexicale et syntaxe).

L'affirmation de Chomsky sur le caractère inné du langage et de son acquisition comme un problème logique plus qu'empirique rencontre alors une forte opposition de la part des linguistes cognitifs et fonctionnels, ainsi que des psycholinguistes (Bates, Bretherton, & Snyder, 1991; Slobin, 1973).

Parmi les principaux travaux en linguistique cognitive, citons entre autres Lakoff et son travail sur la métaphore et la métonymie (Lakoff, 1987) ; Langacker et sa space grammar, puis cognitive grammar (Langacker, 1987) ; Talmy, les force dynamics et le linguistic imaging system (Talmy, 2000) ; Fauconnier et la théorie des espaces mentaux (Fauconnier, 1994).

4 Approches linguistique du temps

Nous venons de présenter quelques approches du traitement du temps dans le langage sans parler de la principale discipline qui s'est penchée sur la question du temps : la linguistique. L'étude du temps dans les langues a évolué avec la linguistique pour passer d'une approche descriptive vers des modèles dynamiques des phénomènes temporels.

Nous présentons dans ce chapitre une liste des principaux phénomènes, en nous concentrant dans une première partie sur l'aspect – qui est le cœur de notre modèle –, puis de manière plus synthétique dans une seconde partie sur la temporalité et la modalité.

4.1 L'aspect

4.1.1 Classes aspectuelles

Vendler (Vendler, 1967) a proposé de catégoriser les verbes et leur compléments (ou **procès**) en classes aspectuelles. Cette catégorisation porte sur l'aktionsart puisque elle concerne (a priori) les valeurs aspectuelles propres à une situation. La classification a souvent été reprise et modifiée, mais nous nous contenterons ici de citer les catégories et leurs traits proposés par Vendler pour donner une introduction à l'aktionsart. Il distingue quatre classes aspectuelles :

- Les **états** (state) sont statiques, duratifs et atéliques

Ex : *be in the house.*

- Les **accomplissements** (accomplishment) sont dynamiques, duratifs et téliques.

Ex : *build a bridge.*

- Les "**achèvements**" (achievement) qui sont dynamiques, instantanés et téliques.

Ex : *realize an error.*

- Les **activités** (activity) qui sont dynamiques, duratifs et atéliques.

Ex : *walk*.

Chez d'autres auteurs, les formulations peuvent changer, tout comme le sens exact des traits. Par exemple, Gosselin (1996) parle aussi d'états, d'accomplissements, d'achèvements et d'activités pour distinguer les classes aspectuelles (qu'il appelle "type de procès"). Deux traits sont identiques : la dynamicité et la ponctualité, tandis que le troisième change : il parle de "bornage" plutôt que de télécité, en précisant qu'il existe un bornage extrinsèque et un bornage intrinsèque (voir R. S. Jackendoff, 1992). Nous reprenons ici la liste de Gosselin¹² et ses exemples :

- Les **états** sont [-dynamique], [-borné], [-ponctuel]

Ex : *être fatigué, savoir quelque chose, aimer le chocolat, etc.*

- Les **accomplissements** sont [+dynamique], [+borné], [-ponctuel].

Ex : *manger une pomme, rentrer chez soi, courir un 100 mètres, etc.*

- Les **achèvements** sont [+dynamique], [+borné], [+ponctuel].

Ex : *apercevoir un avion, sursauter, atteindre un sommet, etc.*

- Les **activités** sont [+dynamique], [-borné], [-ponctuel].

Ex : *marcher, courir, regarder un tableau, etc.*

Si ces quatre classes aspectuelles ont l'avantage de la simplicité, peuvent-elles s'appliquer à toutes les situations ? Est-il même nécessaire de disposer de cette classification et n'est-il pas suffisant d'utiliser les traits pour disposer d'une compréhension aspectuelle d'un énoncé ?

Afin de mieux comprendre l'intérêt de ces traits et les problèmes que leur identification peut poser, nous allons les décrire dans les chapitres suivants. Nous parlerons des traits déjà mentionnés tel que la **dynamicité**, la durativité et la **télécité** ; puis nous introduirons d'autres caractéristiques clés de l'aspectualité (non restreint à l'aktionsart).

4.1.2 Durativité

Qu'est-ce qui distingue *atteindre le sommet* de *gravir le sommet* ? On pourrait dire que le premier énoncé a une durée nulle mais que le second peut prendre un certain temps. Cette caractéristique

¹² Un moins (-) devant le trait signifie que le trait est absent, un plus (+) qu'il est présent.

est appelée la durativité. Celle-ci indique simplement si la situation décrite par le verbe prend du temps. Pour citer Comrie (1976) :

[...] durativity simply refers to the fact that a given situation lasts for a certain period of time (or at least, is conceived of as lasting for a period of time).

Une situation qui n'est pas présentée comme durative est instantanée. Nous insistons sur le fait que ce type de situation n'a pas de durée du tout, et se situe hors de la dimension temporelle. D'après Comrie, le terme de *punctual* est parfois utilisé comme contre-partie à la durativité avec pour exemple typique un verbe comme *cough* (tousser). Mais cet exemple pose problème : *tousser* a une durée et ne peut donc pas être classifié comme *instantaneous*¹³. Comment pourrions-nous alors *tousser pendant deux minutes* ?

"Instantané" semble plus adapté pour décrire une situation sans durée, mais nous préférons parler dans la suite de non-duratif, ou de **durativité nulle**.

Cela entraîne des interprétations particulières si on complexifie un peu les énoncés :

- (34) *Jean a gravi le sommet en deux heures.*
- (35) *Jean a atteint le sommet en deux heures.*
- (36) *Jean a gravi le sommet pendant deux heures.*
- (37) *#Jean a atteint le sommet pendant deux heures.*

Dans l'exemple (34), nous comprenons que Jean a passé deux heures à *graver le sommet*, à escalader ou marcher, et qu'il est arrivé au terme de son ascension.

Dans l'exemple (35), nous comprenons la même chose, qu'il a passé deux heures à grimper ou marcher avant de se retrouver au sommet. Mais à la petite différence que ces deux heures ne décrivent pas *atteindre le sommet* mais ce qui l'a précédé.

L'exemple (37) est incorrect, aucune inférence n'est faite et on comprend bien que *pendant deux heures* entre en conflit avec *atteindre le sommet*, alors que (36) est acceptable. Est-ce uniquement lié à des durativités incompatibles ou d'autres éléments entrent-ils en compte ? Il est trop tôt pour répondre à la question, aussi continuons notre exploration de l'aspect.

4.1.3 Ordre de grandeur

Considérons les durées des situations et des intervalles de temps donnés :

¹³ C'est aussi le cas pour *sursauter*, donné par Gosselin un peu plus haut.

(38) # *Trois minutes après son voyage en Chine.*

(39) *Trois jours après son voyage en Chine.*

L'exemple (38) est incorrect ("bizarre"), on a bien du mal à comprendre le moment où l'on se situe, alors que (39) est parfaitement correct. Le problème de (38) est la mise en correspondance (ici avec la coordination *après*) de périodes de temps de durée très différentes : *trois minutes* et *son voyage en Chine*. Il y a plusieurs ordres de grandeurs de différence ! La prise en compte des **ordres de grandeurs** est une information importante dans le traitement du langage, et particulièrement pour le traitement de l'aspect (voir page 53 le chapitre [4.1.9 Fréquence](#)). L'ordre de grandeur correspond à un ordre de grandeur temporel "typique" d'une situation ou d'un intervalle de temps donné. Cette "magnitude" est utilisée par les humains pour connaître ou inférer la magnitude des situations mises en parallèle.

L'incompatibilité des ordres de grandeurs est variable et va dépendre des ordres de grandeur typiques dont chacun dispose et de son niveau de tolérance. Dans (38), le lecteur pourra donner (consciemment ou pas) le sens de *le moment où je descend de l'avion qui me ramène de Chine à mon voyage en Chine* et la phrase redeviendra acceptable.

4.1.4 Dynamicité

Est-ce que *courir* est une situation dynamique ? Est-ce que *savoir quelque chose* est dynamique ?

D'après Comrie, la plupart des gens répondront oui et non à ces questions, sans précisions sur la définition de 'dynamique'. Ces distinctions intuitives suggèrent leur importance, et tout particulièrement dans des langues comme l'anglais qui a une forme progressive que ne s'applique presque qu'aux situations dynamiques.

Nous reprendrons, ici essentiellement, la première proposition de Comrie (1976) de la dynamicité. Une situation est dite dynamique si des changements ont lieu, sinon elle sera statique. *Savoir* sera une situation **statique**, alors que *courir* sera **dynamique**.

Si les personnes sont généralement cohérentes entre elles lorsqu'on leur demande de catégoriser les situations, des désaccords apparaissent pourtant : *Être debout* n'indique pas explicitement de changement, mais n'exclut en rien des changements. On peut courir tout en restant debout !

Un autre exemple discuté par Comrie est *the oscilloscope is emitting a pure tone at*

300 cycles per second. La situation est apparemment dynamique puisque l'oscilloscope fait quelque chose, mais n'entraîne pas nécessairement de changements.

Comrie propose de préciser les définitions de statique et dynamique :

With a state, unless something happens to change that state, then the state will continue: this applies equally to standing and to knowing. With a dynamic situation, on the other hand, the situation will only continue if it is continually subject to a new input of energy: this applies equally to running and emitting a pure tone [...]. To remain in a state requires no effort, whereas to remain in a dynamic situation does require effort, whether from inside ([...]), or from outside ([...]). Since punctual situations automatically involve a change of state, they are automatically dynamic: there can be no such thing as a punctual state.

Statique et dynamique sont ainsi des propriétés de situations indépendantes de la langue.

4.1.5 Perfectivité et perspective

La définition la plus commune de l'aspect **perfectif** indique que la situation est achevée. Par exemple :

(40) *Jean a dessiné un cercle.*

Comrie (1976) complète cette définition élémentaire en indiquant que la situation est considérée comme un tout ("completed" en anglais, et pas "complete"). D'après lui la perfectivité peut être inférée et non marquée grammaticalement, comme en français :

(41) *Il régna 30 ans¹⁴.*

Le russe, par contre, marque explicitement le perfectif. Mais ici encore son interprétation ne conduit pas systématiquement à l'achèvement de la situation puisqu'il existe un perfectif futur : *ja ub'ju tebja* (*I shall kill you*) ; ou encore un autre exemple en grec ancien : *boulétaï toũto poiésai* (*he wishes to do this*). Les cas s'avèrent complexes et partagés selon la langue, ainsi Comrie cherche une définition de la perfectivité plus générale (Comrie, 1976, p. 21) :

[...] we may consider the view that the perfective represents the action pure and simple, without any additional overtones.

Cette idée de perspective utilisée pour définir l'aspect perfectif suggère que l'on observe alors une

14 Comrie cite aussi la même phrase à l'imparfait *il régnait trente ans* en arguant que celle-ci est aussi perfective, bien que nous pensions que la phrase soit incorrecte (il est nécessaire d'invoquer un contexte très spécifique pour pouvoir "la dire").

situation de "l'extérieur". L'idée d'un point de vue extérieur à la situation se retrouve chez Smith (Smith, 1991, p. 103) :

Sentences with a perfective viewpoint present a situation as a single whole. The span of the perfective includes the initial and final endpoints of the situation: it is closed informationally. [...] Perfective viewpoints present situations as punctual. The impression of punctuality arises from the nature of the perfective presentation.

Celle-ci propose trois types de "viewpoints" : **perfective**, **imperfective** (qui lui-même peut se décomposer en progressive et non-progressive selon des critères pragmatiques) et **neutral**, et parle de end-points pour qualifier les bornes (Smith, 1997) :

Imperfective viewpoint present part of a situation, with no information about its end-points

L'aspect **imperfectif** serait alors l'observation d'une situation depuis "l'intérieur". Mais selon Comrie, les langues se distinguent dans la marque de l'imperfectivité, au point de distinguer différentes catégorisations. L'anglais fournit des exemples d'énoncés qui distinguent l'habituel et le progressif :

(42) *John used to work here.*

(43) *John was working.*

Tandis que le français écrit dispose du passé simple comme temps perfectif, et l'imparfait comme temps imperfectif (voir Patard, 2007 pour une discussion à ce sujet).

Nous reproduisons la table de Comrie 'Classification of aspectual oppositions' :

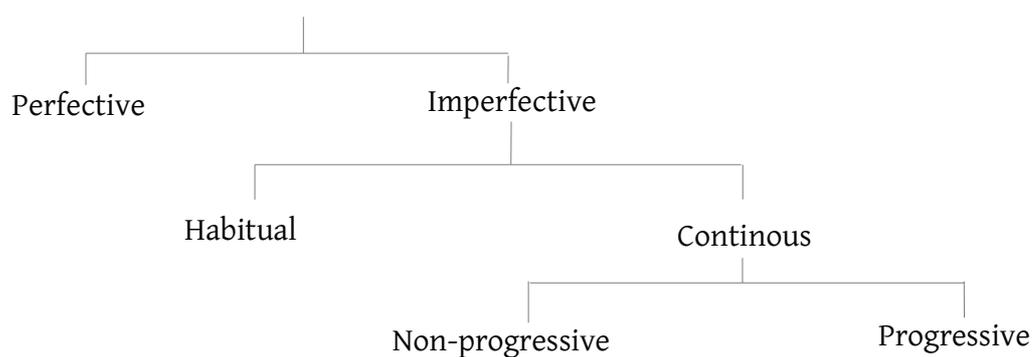


Figure 2 : Classification of aspectual oppositions

Le raffinement de l'imperfectivité ne remet toute fois pas en cause la relation perfectivé/"point de vue extérieur" et imperfectivé/"point de vue intérieur".

Cette idée de perspective se retrouve chez Filip (1999) qui va dans le même sens que Comrie et Smith mais supprime la distinction Habitual/Continuous et se contentant de rattacher progressive et non-progressive à l'imperfectif.

La distinction Non-progressive/Progressive est surtout présente en anglais et indique si situation est statique ou dynamique (respectivement) ; Voir les classes aspectuelles en page 42.

On remarquera que la perfectivité est relative à l'interprétation d'une situation ; elle fait partie de l'aspect "grammatical". Nous insistons sur ce point puisque nous abordons dans le chapitre suivant la notion de télicité, qui elle cherche à catégoriser les situations selon leur structure interne, indépendamment d'un point de vue ou d'une perspective ; que l'on associera à l'aktionsart.

Notons que Talmy (Talmy, 2000) qui parle de perspective, en distinguant continuous et discrete, ne se contente pas d'assigner cette caractéristique aux verbes mais aussi aux noms. En voici une courte représentation (avec ses propres termes) :

Tableau 2 : Plexity et Homogénéité chez Talmy

	Nouns	Verbs
Continuous/Discrete	Mass/count nouns	Action/events

Il est intéressant de voir qu'une seule variable binaire (Continuous/Discrete) puisse être utilisé pour faire des distinctions numériques pour les noms **et** aspectuelles pour les verbes ; très bon exemple à notre avis pour disposer d'un modèle plus parcimonieux.

4.1.6 Télicité

La télicité est une propriété aspectuelle "inhérente" à des classes d'objets lexicaux (Comrie, 1976, p. 41). Elle permet de distinguer les situations qui ont un "accomplissement" de celles qui n'en ont pas. Prenons des exemples :

(44) *Marie dessine un cercle.*

(45) *Marie dessine.*

Supposons que ces deux situations ont la même durée ; La différence entre les deux est que dans

le premier cas Marie vise à atteindre un résultat final, un cercle complet, et que dans le second Marie peut continuer de dessiner encore et encore, sans que l'action ne se termine. (44) ne sera vrai que si le cercle est achevé. L'action peut être interrompue en plein milieu et ne jamais le terminer, mais dans ce cas il n'y a plus d'accomplissement. Si Marie se contente de *dessiner*, qu'elle soit interrompue ou pas ne change rien à la valeur de vérité de l'action.¹⁵ *Dessiner* seul n'a pas de point d'achèvement.

Nous avons parlé précédemment d'objets lexicaux, c'est-à-dire des représentations lexicales de situations. *Dessiner* est une expression atélique et *dessiner un cercle* est une expression télique, alors même que la situation décrite peut être la même. (nous allons revenir sur ce point).

L'identification de la télicité est très importante pour l'aspectualité. Reprenons les exemples (35), (36) et (37)¹⁶:

(35) *Jean a atteint le sommet en deux heures.*

(36) *Jean a gravi le sommet pendant deux heures.*

(37) *#Jean a atteint le sommet pendant deux heures.*

L'exemple (37) n'est pas correct, alors que la seule différence lexicale avec (35) est le complément circonstanciel. Il semble que *atteindre le sommet* impose une contrainte incompatible avec *pendant*, alors que *gravit le sommet* est compatible. Atteindre le sommet et gravir le sommet ne veulent pas dire la même chose, mais surtout ils ne se comportent pas de la même manière. Comment identifier la télicité d'une situation ? Plusieurs approches pour l'anglais ont été proposées et nous reprenons ici celles listés par Borik (2006).

La première approche est celle déjà présentée de la **modification adverbiale**, qui consiste à rajouter *in an hour* pour vérifier si la phrase est télique, ou *for an hour* pour vérifier si la phrase est atélique. Le problème de celle-ci est qu'elle n'est pas systématique et que des interprétations

15 Reprenons les exemples de Klein (1974), mentionnés par Comrie : *Si quelqu'un jouait, et tout en jouant a été interrompu, est-ce qu'il a joué ? Oui, il a joué. [...] Si quelqu'un se noyait, et tout en se noyant a été interrompu, est-ce qu'il s'est noyé ? Non, il ne s'est pas noyé.*

16 Ces exemples sont au passé composé, mais il est possible d'observer les mêmes problèmes à l'imparfait (on commencera les phrases par "ce matin, [...]").

particulières peuvent apparaître, en particulier par l'inchoativité¹⁷, comme dans (35).

La seconde approche est la **conjonction** (Verkuyl 1993), qui utilise des "modifications temporelles".

(46) *Mary drove her car on Monday and on Tuesday* (ambiguous).

(47) *Mary ran a mile on Monday and on Tuesday* (non-ambiguous).

Dans le cas de situations téliques (47) on comprend que deux occurrences différentes et indépendantes ont eu lieu à des intervalles distincts. Dans le cas de situations atéliques (46) il est possible de comprendre l'énoncé différemment : une seule occurrence de la situation aurait eu lieu de manière continue entre les deux bornes proposées (*Monday* et *Tuesday*). Cette interprétation n'est pas possible pour (47).

La troisième approche est le test du **progressif** (progressive test). Il apparaît que la valeur de vérité d'une phrase télique au progressif (on utilisera en français "*en train de*" à la place du progressif anglais) n'est pas la même qu'au simple past. Alors que c'est le cas pour les phrases atéliques. Voici deux exemples :

(48) *Mary was driving the car --> Mary drove the car.*

(49) *Mary was running a mile -/-> Mary ran a mile.*

Quelles sont le ou les points communs entre ces trois approches qui pourrai(en)t permettre de décrire formellement la télicité ?

4.1.7 Point d'achèvement

L'approche la plus commune est celle du point d'achèvement. C'est celle dont nous nous sommes servi dans le paragraphe précédent pour différencier *Marie dessine un cercle* et *Marie dessine*. C'est une notion qui prend plusieurs appellations : telicity (Depraetere, 1995; Filip, 1999; Smith, 1991, 1997), (set) terminal point (Comrie, 1976; Krifka, 1989), boundedness (Declerck, 1979) et culmination (Parsons, 1990).

¹⁷ L'inchoativité indique le début, le commencement d'une situation.

Seulement Borik (2006) soulève plusieurs critiques avec cette approche : il est possible d'utiliser différentes expressions pour décrire la même situation, tout dépend alors de la volonté d'expression du locuteur. La course que Marie a faite hier peut être décrite de plusieurs manières (traduites depuis Borik) :

(50) *Marie courait dans le parc hier.*

(51) *Marie a couru 10 kilomètres dans le parc hier.*

(52) *Marie a couru dans le parc pendant une heure hier.*

La situation est bien terminée (elle a eu lieu hier), néanmoins, (51) la décrit de manière télélique, tandis que (52) la décrit de manière atélique. Ce sont les représentations qui sont téléliques ou non, pas les situations elles-mêmes.

L'utilisation d'un point d'achèvement pose alors problème. Comment parler de point d'achèvement si on s'interdit de décrire des situations ?

La situation décrite par la phrase *Franz Kafka vivait à Prague* prend fin – s'achève – à la mort de Kafka, mais l'expression n'est pas télélique. À partir de cela, chaque situation ayant pris place dans le passé est potentiellement achevée, ce qui ne donne pas pour autant des expressions téléliques.

Depraetere (1995) propose deux stratégies pour identifier des situations achevées. Elle fait la distinction entre inherent (intended) end-point et la temporal boundary d'une situation. Voici ses définitions pour telic et bounded :

A clause is telic if the situation is described as having a natural or an intended endpoint which has to be reached for the situation as it is described in the sentence to be complete and beyond which it cannot continue. Otherwise it is atelic.

A sentence is bounded if it represents a situation as having reached a temporal boundary, irrespective of whether the situation has an intended or inherent endpoint or not.

Mais comme le fait remarquer Borik, ces définitions créent de nouvelles difficultés, comme les illustrent les exemples suivants :

(53) *Sheila deliberately swam for 2 hours* (telic).

(54) *Judith played in the garden for an hour* (bounded).

L'utilisation de *deliberately* pourrait correspondre à la notion de 'intended' donnée par Depraetere. Mais que ce passe-t-il si Sheila a perdu le compte du temps et a nagé deux heures sans

le vouloir expressément ? L'expression deviendrait alors atélique selon sa définition, bien que l'on ait du mal à l'interpréter cette manière.

Une autre proposition est celle de Declerck (1979), qui insiste sur la notion de termination en prenant en compte celle d'homogénéité. Declerck propose qu'un énoncé représente une situation comme télique ou atélique. Avec ses mots 'bounded expressions represent situations as terminating'. À nouveau Borik pointe des failles à cette approche. Prenons un exemple :

(55) *From 1994 till 1996 John lived in New York.*

À première vue, la situation décrite par la phrase est terminée ('terminated'), mais rien ne permet de dire que John ne vit pas toujours à New York. Ainsi, bien que la phrase présente une situation avec des bornes clairement délimitées, la situation *John lived in New York* n'est pas forcément achevée.

Borik mentionne aussi que l'exemple (55) passe les tests du progressif et d'homogénéité et ne peut donc être catégorisé en tant que télique.

On voit que cette notion de télicité, très importante dans la compréhension d'un énoncé est encore vague et critiquable avec l'approche dite du point d'achèvement. Il existe une seconde approche qui est celle de l'homogénéité..

4.1.8 Homogénéité

L'approche de l'homogénéité a déjà été utilisée dans la classification de Vendler, puis reprise et discutée par d'autres (Dowty, 1979; Hinrichs, 1985; Verkuyl, 1996).

Une situation est dite **homogène** si on peut utiliser le même énoncé pour la décrire dans sa totalité ou juste en partie.

Marcher par exemple peut être utilisé pour décrire la situation d'une marche quotidienne de deux heures dans le parc du centre-ville, tout autant que quelques secondes de cette même situation. Borik en donne la formalisation suivante :

(56) P est homogène ssi $\forall x, y (P(x) \& (y \subset x) \rightarrow P(y))$

On notera que cette distinction n'est pas uniquement liée au domaine temporel, mais peut aussi

s'appliquer aux noms (une partie d'un litre d'eau est toujours de l'eau, une partie d'une chaise n'est plus une chaise).

Une autre propriété intéressante liée à l'homogénéité (son "inverse" si on veut) est la cumulativité, qui indique que la somme de même "objets" donne toujours le même objet (du lait additionné avec du lait donne toujours du lait), au contraire des objets non-cumulatifs (une chaise plus une chaise donnent **deux** chaises).

Quine (1960) objecte à la propriété d'homogénéité qu'elle n'est valable qu'à un certain grain : si on descend à un certain niveau de détail, les atomes qui composent une molécule d'eau (ou plus petit encore) seront difficilement perçus comme de l'eau. Et une fraction de seconde de marcher n'est pas vraiment marcher. Ce qui fait que certains auteurs (Krifka, 1989) ont préféré utiliser la notion de cumulativité plutôt que d'homogénéité pour catégoriser la télicité.

L'approche cumulative entraîne un autre problème. Prenons l'activité suivante : *Pousser un chariot*. Les tests de télicité indiquent qu'il est atélique. Mais si on s'imagine quelqu'un pousser un chariot à un endroit, et une seconde personne pousser un autre chariot dans un lieu différent, il devient difficile de considérer le cumule de ces deux situations comme *pousser un chariot*, et on lui préférera *pousser deux chariots*.

4.1.9 Fréquence

Le nombre d'occurrences d'une situation est un autre des attributs de l'aspectualité. Il peut être clairement déterminé comme dans :

(57) *L'an dernier, Marie mangea trois fois à la cantine.*

Ou correspondre à une notion d'habitude, c'est-à-dire un nombre répété de la même situation sans que l'on en connaisse le nombre :

(58) *L'an dernier, Marie mangeait à la cantine.*

(59) *L'an dernier, Marie a mangé à la cantine.*

Remarquons que l'interprétation de (58) est que Marie mangeait de manière régulière à la cantine (mais on en sait pas exactement à quelle fréquence), mais aussi qu'elle n'y mange sans doute plus cette année.

La même interprétation est possible avec (59), en plus d'une seconde interprétation qui rend

possible la situation de *manger à la cantine* unique (ce qui n'est pas possible avec (58), sauf à imaginer un contexte supplémentaire qui vient se placer pendant la situation).

Ces déductions de fréquence et d'habitude semblent directement liées aux temps et à la magnitude (ordre de grandeur) des situations décrites. L'interprétation de l'exemple (60) ne peut plus être habituelle, bien qu'elle puisse encore être répétée (*elle a mangé à la cantine le midi et le soir*), alors que la situation dans (61) sera presque systématiquement unique (on ne diffielement *manger à la cantine* plusieurs fois à midi).

(60) *Hier, Marie a mangé à la cantine.*

(61) *Ce midi, Marie a mangé à la cantine.*

4.1.10 Phasage

Les situations ne sont pas forcément perçues dans leur totalité, nous l'avons vu avec la perfectivité. Il est possible de "montrer" l'ensemble d'une situation ou de celle-ci sans perception des bornes, mais aussi de ne "montrer" qu'une partie de la situation. Les parties discernables sont le début et la fin d'une situation, tandis qu'il est également possible de montrer une partie non-précise. On parle des différentes phases d'une situation.

Ces distinctions sont soit déduites grâce à la pragmatique (voir [4.5 Pragmatique](#), page 65), soit données explicitement avec des verbes ou des prépositions, comme dans les exemples suivants :

(62) *Elle finit ses devoirs.*

(63) *Cette croyance était bien implantée au début du siècle.*

Dans (62), la situation de *faire ses devoirs* est en phase terminale. *Elle* les avait commencés depuis un certain temps et on veut montrer avec le verbe *finir* la fin de la situation *faire ses devoirs*.

Dans (63), la *croyance* n'occupe pas les cents années du *siècle*, mais uniquement une partie : son début. On n'en sait d'ailleurs pas plus, l'étendue exacte de *début du siècle* nous est inconnue.

4.2 Temporalité, modalité et autres éléments pour le traitement du temps

Le sujet principal de cette thèse est l'aspectualité, mais lorsque l'on traite du temps dans les langues il est difficile de parler uniquement d'aspectualité et de ne pas aborder la temporalité et la modalité. Par exemple, en français les temps de conjugaison apportent non seulement une information aspectuelle (différences entre imparfait/passé composé) mais aussi une valeur temporelle (distinction courante entre passé/présent/futur) et modale (certitude/éventualité/possibilité). Nous allons voir dans les parties suivantes quelques points importants pour notre sujet, sans toutefois réaliser un état de l'art complet de la temporalité et de la modalité.

4.2.1 Passé-Présent-Futur

L'un des premiers attributs évoqués lorsque l'on parle du temps en général est le trio "passé – présent – futur" : Une situation est alors souvent décrite comme antérieure au moment présent (c'est un passé), simultanée au moment présent, ou postérieure (c'est un futur). C'est l'approche la plus classique : Elle a besoin d'un "moment" présent (point ou intervalle) qui est souvent décrit comme le moment d'énonciation (speech time) (Hinrichs, 1986; W. Klein, 1994; Reichenbach, 1947b) et d'opérateurs relationnels (avant, pendant, etc.).



Figure 3 : Représentation classique des temps

Reichenbach a une approche minimaliste et n'utilise que les deux opérateurs "précède" et "simultané" pour trois relations possibles (on a A "précède" B, mais aussi B "précède" A). Nous reviendrons plus en détail sur le système de Reichenbach dans le chapitre sur le point de référence (p. 58).

Mais comment ces approches traitent-elles des énoncés utilisant le présent pour exprimer un futur proche ?

(64) *Demain Jean mange à la cantine.*

Le présent est-il capable d'exprimer un présent **et** un futur ? *Demain* est-il assimilé au présent ?

Gosselin (1996) donne au présent français les instructions du "présent" (l'intervalle d'énonciation est simultané à l'intervalle du procès, voir chapitre [5.2 Gosselin](#) page 71 pour plus de détails) et reconnaît alors un **conflit** entre la valeur future du circonstanciel *demain* et la valeur présent du présent simple. Le conflit est **résolu** en appliquant une instruction supplémentaire.

Dans le modèle de Reichenbach (voir page 67), certains temps grammaticaux acceptent plusieurs instructions. Ce n'est pas le cas du simple présent, mais Borik mentionne en citant Verkuyl et Loux-Shuringa (1985) et Vikner (1985) que certains temps grammaticaux de l'anglais peuvent recevoir plusieurs interprétations, par exemple le simple future avec les représentations S,R-E ou S-E,R ou S-R-E.

Un modèle comme celui de Reichenbach est alors contraint d'utiliser ces instructions multiples, mais il est tout à fait possible d'avoir des modèles qui n'ont qu'une seule instruction par temps, mais des mécanismes supplémentaires non liés aux temps qui permettront d'obtenir après un calcul plusieurs interprétations (c'est ce que fait Gosselin, ou le modèle que nous proposons).

Il existe des approches différentes de la représentation passé-présent-futur et distinctes d'un langage particulier. Là où Reichenbach parle de temps relatif (relatif à un moment d'énonciation, de référence), il est possible de représenter le temps de manière "absolue". Par exemple, plusieurs variations où les situations sont caractérisées dans le temps par plusieurs dimensions : la distance, le révolu et l'imaginaire. Dans l'une de ces variations, ces trois dimensions peuvent prendre chacune deux valeurs : proche ou lointain, révolu ou non révolu, imaginaire ou certain. Mettons de côté, pour le moment, l'imaginaire pour voir comment représenter le "passé – présent – futur" avec les deux autres valeurs binaires.

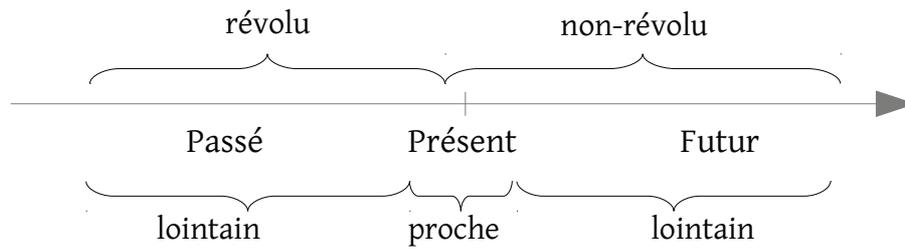


Figure 4 : Les temps avec des valeurs binaires

Dans cette représentation, le présent est un temps proche et non-révolu, le futur est un temps lointain et non-révolu, et le passé est un temps révolu (proche ou lointain, peu importe). En caractérisant le présent comme un temps proche et non révolu, celui-ci ne demande ni valeurs ou instructions supplémentaires, ni résolutions de conflits, pour être compatible avec des énoncés comme (64).

Ce système permet en outre de fournir une explication pour la différence entre le futur simple et le futur périphrastique¹⁸. Prenons deux exemples :

(65) *Marie va aller faire des courses.*

(66) *Marie ira faire des courses.*

La différence entre (65) et (66) n'est pas nette mais serait liée à une notion de proximité : Le futur périphrastique indique un temps futur proche tandis que le futur simple serait plus lointain. On s'en rend mieux compte en leur ajoutant un circonstanciel temporel distant, par exemple *l'année prochaine*. Dans ce cas (66) nous semble plus naturel.

Bien entendu, des langues comme le français ou l'anglais disposent de temps grammaticaux dit "complexes" (comme le plus-que-parfait) qui ne se satisfont pas de cette approche et nécessitent des informations ou instructions supplémentaires afin d'indiquer par exemple l'antériorité à une situation elle-même passée. C'est là que les approches inspirées de Reichenbach expriment toute leur force.

¹⁸ Également appelé "futur proche". C'est le futur qui fait appel au semi-auxiliaire "aller", comme dans l'exemple (65)

4.2.2 Moment de référence

Le concept de moment de référence est bien connue en linguistique. Il a été introduit par Reichenbach (1947a) qui cherchait à modéliser les temps "complexes" de l'anglais. Reichenbach propose de distinguer trois "points" pour rendre compte du système de temps anglais : **S**, le moment d'énonciation (speech time ou point of speech), **E** le moment de "l'événement" (point of event), et enfin **R** le moment de référence (point of reference).

Seulement Reichenbach ne propose pas de définition formelle pour ces trois points, mais uniquement des exemples linguistiques, ce qui laisse la voie aux interprétations. En ce qui concerne S et E, les avis sont plutôt concordants : S est le moment (point in time) où est formulé l'énoncé à propos d'un l'événement particulier, et E est le moment où prend effectivement place cet événement. Il suffit ensuite de disposer d'opérateurs relationnels entre ces deux points pour en déduire le temps (tense : passé, présent, futur) de l'événement : Si E précède S ($E < S$), c'est un passé, si E est simultanée à S ($E = S$) c'est un présent, et si E est antérieure à S ($S < E$), c'est un futur. Par exemple :

(67) *Elle voit Jean.* $S = E$

(68) *Elle voyait Jean.* $E < S$

(69) *Elle verra Jean.* $S < E$

Le problème avec seulement deux points, c'est qu'il n'est pas possible de représenter les temps dit "complexes" (qui correspondent aux perfect tenses en anglais), comme le present perfect, et de les distinguer des temps "simples", comme le simple past. C'est là qu'intervient le point de référence. Celui-ci va insister sur le moment que l'on souhaite "montrer". Par exemple :

(70) *Mary has seen John* $E < R < S$

(71) *Mary saw John* $E < R = S$

Dans (70) le point de référence R est simultanée à S et postérieure à E, indiquant que l'énoncé présente ("montrer") l'état fini de E. Dans (71) le point de référence R est simultanée à E, indiquant que l'élément pertinent est l'action elle-même et pas l'état résultant. Le past perfect est alors représenté par la relation $E < R < S$:

(72) *Mary had seen John* $E < R < S$

Dans des phrases à circonstanciels, comme *I had met him yesterday*, *yesterday* correspondrait selon Reichenbach au point de référence. Mais comment prendre en compte et différencier des circonstanciels comme *in 10 min* et *during 10 min* ? D'autres problèmes ont été mentionnés sur les lacunes du modèle de Reichenbach (par exemple le cas des *embedded clauses*, (Hwang & Schubert, 1992).

Il faut remarquer que le système de Reichenbach est fait pour traiter des temps anglais et qu'il ne prétend pas répondre à toutes les questions de temporalité dans les langues. Le concept de point de référence garde cependant toute sa pertinence et la plupart des travaux postérieurs à Reichenbach vont le réutiliser et le formaliser (voir à ce propos Gosselin et Passonneau 1996; 1988). Comme le disent Kamp et Reyle (1993, p. 594, Part II) : "We will follow Reichenbach in what we consider to be the spirit of his proposal, though not in details".

On notera que ce point de référence, qui permet de "montrer" ce qui doit être vu peut rendre compte de la perfectivité d'un énoncé et se rapproche ainsi du concept de point de vue : "Ce que l'on montre" avec le point de référence et "la position de l'observateur (à l'intérieur ou à l'extérieur) par rapport à une situation" partage une analogie avec le regard.

Gosselin (1996), qui travaille sur le français reprend l'idée du point de référence, mais en transformant les **points** en **intervalles**¹⁹, et en ajoutant des intervalles circonstanciels au lieu de limiter les compléments circonstanciels (*hier, pendant dix minutes*) à l'intervalle de référence. L'intervalle d'énonciation est noté [01,02], l'intervalle du **procès** (qui correspond au point of event de Reichenbach) est noté [B1,B2], l'intervalle de référence est noté [I,II], et l'intervalle circonstanciel est noté [ct1,ct2]. Voici l'un des exemples qu'il donne :

(73) *Samedi dernier, Luc a été à la pêche.*

Voici la représentation graphique de (73) donnée par Gosselin :

19 Chez Gosselin, un intervalle correspond à la portion de l'axe temporel qui est occupée par une situation, et est délimité par un couple de points (les bornes).

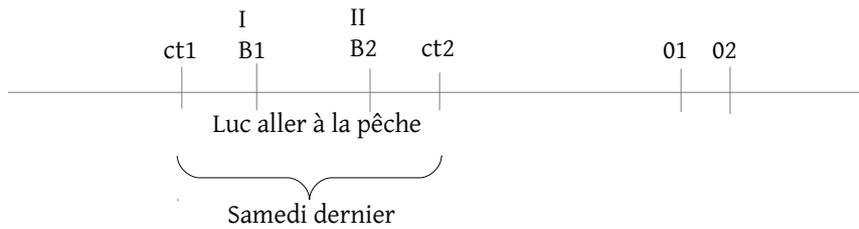


Figure 5 : Représentation graphique chez Gosselin

Reinhart (1986, 2000) développe aussi le concept de temps de référence²⁰. Tout comme Gosselin, elle transforme les points de Reichenbach en intervalles, et propose un système de relations particulières :

- a. the E-R relation is fixed, i.e. $E \subseteq R$ by default (except for progressive);
- b. the S-E relation determines the truth conditions and the temporal interpretation of a sentence.
- c. the S-R relation determines perspective and morphological tense.

4.2.3 Opérateurs relationnels

Nous avons vu dans les parties précédentes que la temporalité peut s'exprimer avec des "temps" (d'énonciation S, de référence R, de l'événement E) et des relations entre ces temps. Il faut utiliser des opérateurs binaires pour définir les relations temporelles entre deux situations (nous rappelons que selon notre définition, les **situations** englobent aussi les expressions de durées, comme *dix minutes*). Que les deux éléments mis en relation soient des points ou des intervalles, il est toujours nécessaire d'en avoir deux. Un énoncé "complet" utilisant l'adverbe *après* aura cette forme :

(74) *Marie est partie après le match.*

ou encore

(75) *Marie est partie après dix minutes.*

Les modèles de type Reichenbach²¹ ont besoin d'opérateurs binaires, des expressions comme

²⁰ Ne disposant des sources originales de Reinhart, nous reprenons ici essentiellement l'analyse que Borik (2006) en fait.

²¹ C'est-à-dire qui exprime la temporalité avec des relations binaires entre le temps d'énonciation, le temps de l'événement, le temps de référence, et éventuellement d'autres temps.

"before S" étant incomplètes (Il est nécessaire d'avoir deux variables : "X before Y") : On ne sait pas si c'est E ou R qui est avant S.

(76) *après le début du match*

(76) est pourtant compréhensible. Bien sur, il sera prononcé ou écrit dans un contexte particulier, et on pourrait argumenter que si (76) est compréhensible c'est parce que l'on dispose des moyens, du contexte, pour combler les vides, c'est-à-dire répondre à "que se passe-t-il après le match ?". Pourtant, même si (76) n'est pas un énoncé complet, il est possible de comprendre quelque chose, et peut-être même de deviner le contexte ; c'est un indice que l'interlocuteur aura bien analysé ce bout de phrase sans attendre la suite ou de disposer un énoncé complet.

Considérons pour le moment que les opérateurs relationnels sont des éléments impératifs pour le traitement du temps et que la plupart des auteurs les traitent comme des opérateurs binaires. Nous verrons plus tard (Chapitre [6.3.3 L'ordonnement](#)) que cela peut poser des problèmes selon l'approche choisie pour le traitement du temps.

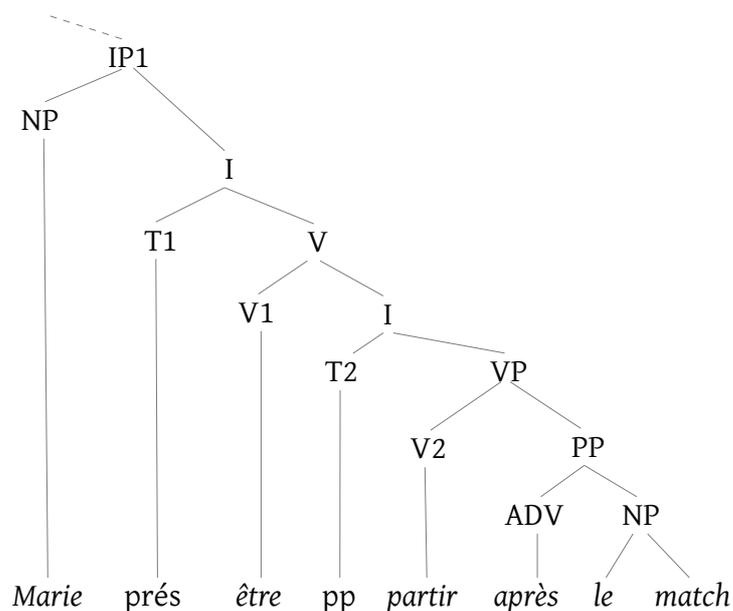


Figure 6 : Représentation syntaxique des informations temporelles

4.2.4 Détermination

Les éléments ayant un impact sur l'interprétation temporelle d'un énoncé ne sont pas toujours aussi évidents que les temps de conjugaison ou les adverbes temporels. Par exemple :

(77) *Marie parlera pendant 45 minutes.*

(78) *Marie parlera pendant ces 45 minutes.*

La première interprétation de (77) et de (78) est similaire : Marie parlera pendant une durée de 45 minutes, peut-être un peu moins ou un peu plus, mais certainement pas pendant 10 minutes.

(78) a une seconde interprétation que (77) n'a pas et ne peut pas avoir : Marie prendra la parole au cours de la période de 45 minutes. Elle pourra dire un mot ou parler plus longtemps, on ne sait pas. Cette seconde interprétation est rendu possible grâce au déterminant *ces*. Nous ne donnerons pas d'explication pour le moment et nous contenterons de faire remarquer l'importance du déterminant dans l'interprétation temporelle. À noter qu'il est possible de rendre acceptables certaines phrases incorrectes. Si on reprend l'exemple (37) :

(37) *#Jean a atteint le sommet pendant deux heures.*

qui devient

(79) *Jean a atteint le sommet pendant ces deux heures.*

Ce mécanisme est connu dans la littérature sous le terme d'*anchoring* et correspond aussi aux énoncés avec une date. Par exemple :

(80) *Jean a atteint le sommet le vendredi 7 décembre.*

(81) *Hier, Jean a atteint le sommet.*

Dans les deux cas, l'atteinte du sommet est "ancrée" dans le temps par les circonstanciels temporels. La méconnaissance exacte de la situation (l'année dans (80), l'heure et les minutes dans (81), par exemple) n'a pas de conséquences sur l'interprétation des énoncés, pourvu que la situation soit déterminée.

Nous ne rentrerons dans les détails de cette détermination temporelle que dans le chapitre sur le modèle (page 123), et laissons pour le moment de côté certaines ambiguïtés (Des énoncés comme *en hiver* ou *le mardi* sont-ils ancrés ?).

4.3 Modalité

Nous avons rapidement mentionné la modalité dans le chapitre sur le passé-présent-futur,

lorsque nous parlions "d'imaginaire". Et n'en parlerons que superficiellement dans ce chapitre. C'est un domaine que nous n'avons que peu abordé lors de la thèse mais que nous mentionnons néanmoins parce que faisant partie intégrante du traitement du temps en général. Qu'est ce que la modalité ?

(66) *Marie ira faire des courses.*

(82) *Marie irait faire des courses.*

Dans ces exemples, les situations se situent dans le futur (proche ou lointain on ne sait pas), et leur aspectualité est similaire. Mais pour (66) on est quasiment certain que la situation aura lieu, alors que c'est beaucoup plus incertain pour (82). La modalité correspond à ce domaine du possible, du certain, de l'imaginaire. Elle peut aussi être exprimée avec des adjectifs, comme *sûrement*²², *probablement*, *peut-être*, etc., ou des articles, comme *si*.

La modalité en français peut avoir un impact direct sur les temps utilisés et peut même aller à l'encontre des valeurs "traditionnelles" temporelles. L'exemple le plus classique étant l'imparfait :

(83) *Si elle venait [...]*

La situation dans (83) n'est certainement plus un passé, mais plutôt un futur proche hypothétique. A noter l'impact de formulation de l'incertitude sur ce genre de construction :

(84) *Il ne sait pas si elle vient (viendra) [...]*

On signalera des travaux récents de Saurí et Pustejovski (2012) soucieux de parcimonie qui parlent de factuality plutôt de que modality.

4.4 Attitude propositionnelle

Nous avons observé lors de nos recherches un phénomène intéressant lié aux notions "d'exceptionnalité" et de négation : Certains énoncés a priori incorrects se révélaient acceptables à condition de considérer des situations exceptionnelles ou "déterminées" (comme un record ou une interdiction non-respecté).

Comparons ces deux exemples :

(85) *Marie est allée en Italie l'année dernière.*

22 Qui indique paradoxalement que l'on n'est pas sûr !

(86) *Marie a bu du champagne l'année dernière.*

Il nous semble que l'énoncé (86) implique que Marie a brisé une interdiction, ou qu'elle a bu du champagne pour la première fois. Tandis que cette "exceptionnalité" n'apparaît pas dans (85). Remarquons que la différence aspectuelle entre ces deux énoncés est la télicité de la situation (*aller en Italie* est télique, *boire du champagne* est atélique). Mais la disparition de la dimension temporelle peut aussi apparaître avec la modification du circonstanciel de temps, sans lien avec la télicité :

(87) *Marie a bu du champagne pendant deux heures.*

L'ajout de *pendant deux heures* n'implique plus comme dans l'exemple (86) une "exceptionnalité".

Prenons un autre exemple :

(88) *Marie a lu la lettre pendant l'entracte.*

La première interprétation indique que la lecture de la lettre c'est étirée sur la totalité de l'entracte, et qu'il est même possible que la lecture n'est pas complète.

La seconde interprétation implique une "attitude propositionnelle" sur *lire la lettre* (on pourra changer de ton sur cette partie à l'oral). Dans ce cas, l'étalement de la situation avec *l'entracte* n'est plus pertinent : *la lecture de la lettre* a eu lieu à un moment donné de l'entracte, qu'importe si ça a été pendant une partie ou la totalité : Nous ne serions plus dans le domaine temporel²³.

(89) *Marie a bu du champagne l'année dernière.*

Il nous semble qu'en imaginant quelqu'un prononcer (89), il prendrait un air surpris ou outré, comme si (89) avait un point d'exclamation ou d'interrogation. Marie buvait du champagne l'année dernière, mais n'en boit pas en temps normal. Ou elle en buvait alors qu'elle n'en avait pas le droit : On s'attendait au contraire.

Dans l'exemple (89), la situation exprimée par l'élément "prédié" est mise en relief²⁴. (88) et (89) expriment des exceptions par rapport aux attentes d'un locuteur.

Attention, on parlera ici de prédication principale, et pas uniquement de prédication : Il est en effet possible de prédiéquer n'importe quel élément de l'énoncé si c'est une réponse ("en écho"),

²³ Nous en faisons là l'hypothèse. Celle-ci nous servira plus tard dans la description de notre modèle.

²⁴ Son contraire est considéré comme "normal" : *Marie n'a pas bu de champagne l'année dernière.*

par exemple :

- (90) - "Jean a bu du champagne l'année dernière."
- "Non, **Marie** a bu du champagne l'année dernière."

La prédication principale correspond à l'information (nouvelle) que le locuteur veut introduire dans la conversation.

Mais pourquoi parlons-nous de ce phénomène ? Il a effectivement un impact sur l'interprétation temporelle des énoncés, mais les modèles temporels récents (Gosselin, 1996; Kamp & Reyle, 1993; de Saussure, 2003; Schilder & Habel, 2001) y font référence sans le formaliser et renvoient son traitement à la pragmatique. Nous verrons dans le chapitre 7.3.3.2 Focus-out et "prédication" temporelle qu'il est possible de prédire ce phénomène de prédication temporelle avec un mécanisme simple et sans faire appel à la pragmatique ou la pertinence.

Attention, cela ne signifie pas que les connaissances générales et du contexte n'ont pas leur mot à dire dans les interprétations temporelles !

4.5 Pragmatique

Voici une définition de la pragmatique d'après Mey (2001) :

Pragmatics is a subfield of linguistics which studies the ways in which context contributes to meaning. Pragmatics encompasses speech act theory, conversational implicature, talk in interaction and other approaches to language behavior in philosophy, sociology, and linguistics and anthropology. Unlike semantics, which examines meaning that is conventional or "coded" in a given language, pragmatics studies how the transmission of meaning depends not only on structural and linguistic knowledge (e.g., grammar, lexicon, etc.) of the speaker and listener, but also on the context of the utterance, any preexisting knowledge about those involved, the inferred intent of the speaker, and other factors. In this respect, pragmatics explains how language users are able to overcome apparent ambiguity, since meaning relies on the manner, place, time etc. of an utterance.

Comment intervient-elle dans le traitement de la temporalité ? Voyons pour cela quelques exemples discutés par Moens et Steedman (1988) et proposés par Ritchie (1979) :

- (91) *When they built the 39th Street bridge ...*

- a. ... a local architect drew up the plans.
- b. ... they used the best materials.
- c. ... they solved most of their traffic problems.

L'ordonnement de la clause principale (a, b ou c) par rapport à la clause secondaire introduite par *when* est différent à chaque fois : Dans (a) la construction du pont est postérieure aux plans (c'en est une conséquence), dans (b) les deux clauses sont simultanées dans le temps, et enfin dans (c) la construction du pont est antérieure à la résolution des problèmes de trafic (c'en est la cause).

Nous parlons bien ici de causes et conséquences. Les différents sens ne sont pas les résultats d'une supposée ambiguïté de *when* mais les interprétations faites à partir de notre connaissance du monde et des relations de causes conséquences. Pour se convaincre du caractère "contingent" ("contingency") de *when*, Moens et Steedman proposent l'énoncé suivant :

(92) # *When my car broke down, the sun set.*

(92) n'est pas correct, parce qu'il n'existe pas de causalité ou de lien logique entre les deux situations (sauf dans un monde imaginaire et fantastique où ce serait le cas...).

When n'est pas une exception, et il est possible de mettre en évidence l'importance de la pragmatique dans l'ordonnement des situations avec d'autres exemples :

(7) *Marie est tombé. Jean l'a poussé.*

On considère que Marie tombe parce que Jean l'a poussée. Il est néanmoins possible que Marie tombe d'abord, puis que Jean la pousse ensuite. C'est d'ailleurs l'ordre que l'on préférera si aucune information pragmatique ne nous dit le contraire.

Avant de conclure cette partie, revenons à l'exemple (91) : Nous avons mentionné dans le chapitre sur le moment de référence que l'on peut "montrer" une partie seulement de la situation décrite. C'est ce que Moens et Steedman illustrent ici : Pour (91)(a) on ne présente pas la totalité de la construction du pont, mais son état préparatoire ou le début de la situation. La clause principale est "le moment de référence" au sens de Reichenbach.

5 Des modélisations pour le traitement de l'aspect

Nous présentons dans ce chapitre quelques modèles du traitement de l'aspect – certains traitant aussi le temps – et proposons quelques remarques et critiques par rapport aux critères de parcimonie, de plausibilité cognitive et s'ils donnent les résultats attendus. La liste des modèles choisis n'est de loin pas exhaustive. On se servira dans ce chapitre des critères de parcimonie et de plausibilité cognitives soit pour critiquer un modèle, soit pour s'en inspirer.

Nous ne parlerons pas des modèles statistiques au prétexte que nous les considérons comme peu plausibles cognitivement et peu parcimonieux en raison de l'importante quantité de données nécessaires. De plus ceux-ci ne sont pas capables de repérer les interprétations aspectuelles, comme la répétition dans l'exemple (58).

(58) *L'an dernier, Marie mangeait à la cantine.*

Surtout, nous sommes intéressés de découvrir comment les individus calculent l'aspect, pas d'avoir une liste de toutes les traductions possibles des énoncés !

Ghadakpour (2003), qui n'a pas proposée de modèle complet du temps ou de l'aspect, sera abordée dans le chapitre suivant lorsque l'on introduira le module perceptif et ses interactions avec le module du traitement de l'aspect (aspect "grammatical").

5.1 Reichenbach

Reichenbach publie en 1947 *Element of Symbolic Logic* dans lequel il introduit un modèle de traitement du temps de l'anglais qui influencera durablement les travaux futurs sur le sujet. Le

modèle se distingue (à notre avis) de deux manières remarquables :

La première est qu'il s'inspire directement de la logique et est **minimaliste** par rapport aux nombreux travaux suivants sur le temps (ce qui est explicable puisque ce sont pour la plupart des modèles enrichis, plus complets mais qui traitent davantage de cas). Il n'a besoin que de trois "**moments**" (points dans l'original) et de deux **relations** possibles pour ordonner ces moments entre eux : la simultanéité, que nous noterons ici par le symbole égal "=", et la précédence que nous noterons par un inférieur "<".

La seconde distinction est l'introduction d'un moment de **référence**. En effet, Reichenbach fait intervenir trois moments pour décrire les temps anglais : Le moment d'élocution **S** (ou time of Speech), le moment de l'événement **E** (point of Event) et ce moment de référence **R** (point of Reference). Nous en discutons plus en long dans le chapitre sur le Moment de référence, page 58.

Comment ces éléments sont-ils organisés pour rendre compte des nombreux temps grammaticaux de la langue anglais ?

La relation entre le moment d'élocution S et le moment de l'événement E va déterminer si l'événement se situe dans le passé, le présent, ou le futur. Par exemple :

(93) *She writes a letter.*

(94) *She wrote a letter.*

(95) *She will write a letter.*

On note alors l'ordonnancement des moments de cette manière : (93) S=E (94) E<S (95) S<E. Ce qui correspond respectivement au simple present, simple past, et simple future.

Seulement ces deux moments ne sont pas suffisants pour rendre compte des temps dits "complexes", comme le present perfect. C'est là qu'intervient le moment de référence.

(96) *She has written a letter.*

Le present perfect aura alors cette représentation : (96) E<R=S. Tandis que le simple past sera représenté ainsi : (94) E=R<S.

Dans le premier cas la référence est simultanée à l'événement, dans le second cas, la référence est simultanée à l'énonciation. On peut interpréter cette représentation en considérant que le point de référence permet de montrer ce qui est important²⁵ : Le résultat de l'événement dans (96), et

25 Gosselin (1996) écrit :

... qui représente ce qui est perçu/montré sur l'axe temporel.

l'événement lui-même dans (94).

Voici un tableau des temps anglais avec leur représentation dans le modèle de Reichenbach, tiré de (Mani, Pustejovsky, & Gaizauskas, 2005) :

Tableau 3 : Correspondance entre temps grammaticaux et les relations de Reichenbach

Relation	Reichenbach's Tense Name	English Tense Name	Example
E<R<S	Anterior past	Past perfect	<i>I had slept</i>
E=R<S	Simple past	Simple past	<i>I slept</i>
R<E<S	Posterior past		<i>I expected that...</i>
R<S=E			<i>I would sleep</i>
R<S<E			
E<S=R	Anterior present	Present perfect	<i>I have slept</i>
S=R=E	Simple present	Simple present	<i>I sleep</i>
S=R<E	Posterior present	Simple future	<i>I will sleep (Je vais dormir)</i>
S<E<R	Anterior future	Future perfect	<i>I will have slept</i>
S=E<R			
E<S<R			
S<R=E	Simple future	Simple future	<i>I will sleep (Je dormirai)</i>
S<R<E	Posterior future		<i>I shall be going to sleep</i>

Le système de Reichenbach apparaît comme simple et puissant. Trop expressif même puisque Reichenbach note que le nombre de possibilités va au-delà du nombre des temps anglais, ce qui le conduit à limiter les relations possibles en fixant des contraintes pour le point de référence, réduisant le nombre de possibilités à 9.

The position of R relative to S is indicated by the words 'past', 'present' and 'future'.
 The position of E relative to R is indicated by the words 'anterior', 'simple' and 'posterior', the word 'simple' being used for the coincidence of R and E
 (Reichenbach, 1947a, p. 297)

On constate que seulement trois attributs (E, S et R) et deux relations binaires (simultanéité et précédence) sont suffisants pour rendre compte des temps anglais.

Un unique algorithme est nécessaire pour réaliser ces représentations et il n'est pas nécessaire "d'attacher" de procédures quelconques à un élément du lexique. Malgré ce support au

minimalisme, le modèle de Reichenbach n'est pas exempt de critiques. Il ne parvient pas par exemple à rendre compte de tout les temps grammaticaux d'autres langues (Comrie, 1985; Dowty, 1979) ou échoue à rendre compte des clauses imbriquées (Hwang & Schubert, 1992).

Considérant par exemple l'énoncé suivant :

- (97) *John will find this note when he gets home.*
He will think_(a) Mary has left_(b)

La seconde phrase obtient avec l'analyse Reichenbach la représentation suivante : $E_b < S$ et $R_b < R_a = E_a$

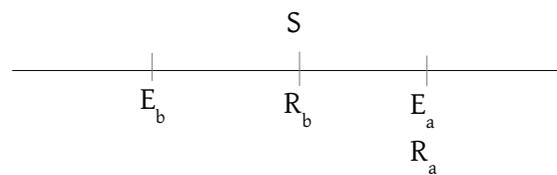


Figure 7 : Exemple (97) selon Reichenbach

C'est-à-dire que John va penser que le départ de Mary (E_b) a pris place avant le moment de l'énonciation (S). Or, l'interprétation n'est pas correcte, puisque John va penser que le départ de Mary n'est que antérieur à sa pensée, c'est-à-dire $S < R_a$ et $E_b < R_b = R_a$, que Hwang et Schubert illustrent de cette manière :

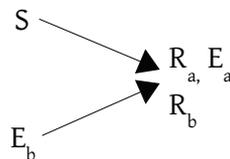


Figure 8 : Exemple (97) selon Hwang et Schubert

De plus, le minimalisme de Reichenbach est aussi possible par son attention sur la temporalité, alors qu'il ne traite pas d'aspectualité (progressive) ou de modalité comme le font les modèles que nous aborderons plus bas.

Malgré des lacunes, le modèle de Reichenbach est suffisamment innovant et pertinent pour orienter la plupart des travaux ultérieurs. Une partie de ces travaux proposera des amendements à l'original, tandis que d'autres n'hésiteront pas à s'inspirer plus librement sans conserver tous les éléments du modèle de Reichenbach.

Nous retiendrons pour notre part qu'il est possible de traiter du temps, de manière encore imparfaite, en ayant un modèle minimaliste : Le modèle de Reichenbach utilise uniquement deux types de relations et trois types de moments. Notre idée est qu'il est possible de faire la même chose – être minimaliste, ou **parcimonieux** dans notre vocabulaire – pour l'aspect.

5.2 Gosselin

Le modèle de Gosselin (1996) entre dans cette catégorie de modèles qui s'inspire de celui de Reichenbach. Celui-ci se concentre sur le français plutôt que l'anglais et ne se contente pas de traiter de temporalité mais aussi d'aspectualité et de modalité (Gosselin, 2005).

Les trois notions que sont *reference*, *event* et *time of speech* chez Reichenbach apparaissent chez Gosselin sous les termes de *référence*, *procès* et *énonciation*, mais ceux-ci ne sont plus des moments (points) mais des intervalles. L'utilisation d'intervalles est assez courante dans les modèles du traitement du temps et permettent entre autre de traiter de l'aspectualité : Il est possible de représenter grâce à eux des débuts, des fins, des chevauchements, des inclusions, etc. Il existe de plus un quatrième type d'intervalle, les intervalles **circonstanciels** qui correspondent aux compléments circonstanciels de temps (*samedi dernier* par exemple).

L'intervalle du procès est noté $[B1, B2]$ ($B1$ et $B2$ étant les bornes de début et de fin), l'intervalle d'énonciation est noté $[01, 02]$, l'intervalle de référence est noté $[I, II]$, et l'intervalle circonstanciel est noté $[ct1, ct2]$.

Gosselin réutilise directement les classes aspectuelles de Vendler : états, activités, accomplissements et achèvements. Nous ne nous étendons pas ici dessus puisque nous l'abordons page 42 dans le chapitre Classes aspectuelles. Mais on peut néanmoins indiquer que ces classes aspectuelles vont avoir un impact sur la représentation des intervalles : Un intervalle est composé de deux bornes (début et fin). Celles-ci peuvent être de deux types différents pour les procès : "intrinsèques", c'est-à-dire lorsqu'elles sont impliquées par le procès ; ou bornes "extrinsèques" sinon.

Nous reproduisons ici son tableau récapitulatif (Gosselin, 1996, p. 56) :

Tableau 4 : Types de procès chez Gosselin

Type de procès	figure	Configuration
État nécessaire	Absence de bornes	États (absence de changements)
État contingent	Bornes extrinsèques	États (absence de changements)
Activité	Bornes extrinsèques	Séries de changements
Accomplissement	Bornes intrinsèques	Séries de changements
Achèvements	Bornes intrinsèques	Changements atomiques

À noter que c'est avec cette catégorisation entre bornes intrinsèques et extrinsèques que le modèle permet de différencier les syntagme *en dix minutes* et *pendant dix minutes* : Le premier est intrinsèque et indique que la durée s'applique à l'intervalle du procès (*il a mangé le gâteau en dix minutes* indique que le procès de manger a une durée de dix minutes). Le second est extrinsèque et indique que le procès a lieu pendant cet intervalle mais qu'il peut se poursuivre plus tard (*il a mangé le gâteau pendant dix minutes, sans le terminer*).

Les moments ayant évolué en intervalles, les relations de Reichenbach ne sont plus suffisantes. Gosselin utilise trois relations entre **bornes** : la simultanéité (notée par un égal "="), la précedence "infiniment proche" (" α ")²⁶, et la précedence mais sans voisinage immédiat (" \downarrow "). "<" englobera les relations " α " et " \downarrow ", tandis que " \leq " englobera les trois relations.

De ces relations entre bornes il est possible d'obtenir (entre autres) des relations entre intervalles comme l'antériorité (tel que $[i,j]ANT[k,l]$ si $j < k$) :

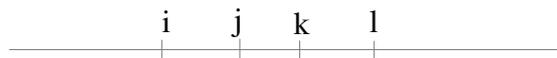


Figure 9 : Antériorité ; ANT

ou la simultanéité (tel que $[i,j]SIMUL[k,l]$ si $(i \leq l) \& (k \leq j)$) :

26 Gosselin justifie l'existence de cette relation par la nécessité de rendre compte de procès "ponctuels" :

[...] qui occupent une durée infiniment petite et référentiellement négligeable (i.e. Qui ne peut être indiquée au moyen d'un circonstanciel de durée) ; ex. : l'étincelle *jaillit* ($B1 \alpha B2$).

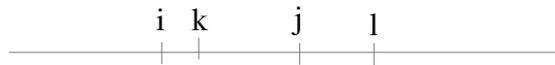


Figure 10 : Simultanéité ; SIMUL

Voici les définitions de la coïncidence et de l'accessibilité (que nous utilisons plus bas).

Coïncidence : $[i,j]$ CO $[k,l]$ si $(i = k) \ \& \ (j = l)$



Figure 11 : Coïncidence ; CO

Accessibilité : $[i,j]$ ACCESS $[k,l]$ si $(i \leq k) \ \& \ (l \leq j)$



Figure 12 : Accessibilité ; ACCESS

Chez Gosselin, un énoncé est analysé comme une séquence linguistique d'**instructions** pour la construction d'éléments **de représentation**. Voici ce qui est dit à propos de ces instructions (Gosselin, 1996, p. 48) :

Les instructions sont codées par les marqueurs linguistiques. La distinction grammaticale entre **lexèmes** et **marqueurs grammaticaux** (morphèmes et constructions syntaxiques) s'avère pertinente dans ce cadre : les instructions associées aux marqueurs grammaticaux portent directement sur la construction d'éléments de représentation (elles peuvent aussi établir des relations entre des éléments construits par ailleurs); tandis que les lexèmes déclenchent l'activation d'un **prototype** associé (le prototype, ou meilleur exemplaire de la classe, constituant le point de repère par rapport auquel se définit, en termes de degrés de similarité, l'opération de catégorisation). Et c'est par l'intermédiaire de ce prototype que des éléments de représentation sont construits dans ER. Comme le prototype inclut à la fois

des propriétés nécessaires et des propriétés contingentes (par exemple, le fait de voler pour un oiseau), il suit que les représentations ainsi construites comportent une part de contingence, et peuvent être au moins partiellement remises en cause par les données contextuelles.

Nous retiendrons qu'il existe des prototypes en plus des éléments de représentation. (les intervalles et leurs relations), et que la construction des éléments de représentation est provoquée par des instructions attachées à des éléments du lexique. Prenons l'exemple de **en + durée** et **pendant + durée** pour illustrer le fonctionnement de l'une de ces instructions (Gosselin, 1996, p. 58) :

- a) construction d'un intervalle [ct1,ct2] (circonstanciel temporel)
- b) ct1 $\{$ ct2 (circonstanciel non ponctuel)
- c) [ct1,ct2] CO [B1,B2] (relation circonstancielle de coïncidence avec le procès)
- d) [I,II] ACCESS [B1,B2] (puisque le circonstanciel porte sur [B1,B2], les bornes du procès doivent être **accessibles** à partir de l'intervalle de référence ; seul l'aspect aoristique (perfectif) est directement compatible avec cette exigence).

Avec comme distinction entre *en* et *pendant* (Be représentant une borne extrinsèque et Bi une borne intrinsèque) :

- e) [*pendant* + durée] : Be1 $\{$ Be2
- f) [*en* + durée] : Bi1 $\{$ Bi2.

Sans rentrer dans les détails de ces instructions, on constate qu'un mot comme *en* peut disposer d'un nombre non-limité d'étapes. Et cela ne concerne que le syntagme *en* + durée ! Pour *en 2012*, *en* aura donc une autre série d'instructions spécifiques.

Ainsi chacun des marqueurs grammaticaux (*en*, *pendant*, *après*, *avant*, *au cours de*, *durant*, etc., plus tout les temps de conjugaisons) et potentiellement certains lexèmes (*hier*, *demain*, *matin/matinée*) auront à leur tour leur propre liste d'instructions. Si on se limite aux marqueurs aspectuo-temporels²⁷, cela correspond à un nombre considérable, bien que fini²⁸, d'instructions.

²⁷ Voici une liste simplifiée donnée par Gosselin (1996, p. 24) : Les verbes et leurs compléments, les temps morphologiques, les types de circonstanciels temporels et leur construction syntaxique, les adverbes et locutions adverbiales d'aspect (itératifs, numéraux), les constructions de phrases complexes, les enchaînements d'énoncés dans le texte et la structure des paragraphes.

²⁸ Même si de nouveaux mots apparaissent régulièrement dans les langues, il est beaucoup plus rare de voir apparaître de nouveaux marqueurs grammaticaux. En percevant le temps de manière topologique, il est attendu

En d'autres termes, le modèle de Gosselin est contraint de lier des algorithmes aux grammèmes et à certains lexèmes. Se pose alors le problème du minimalisme : par rapport à un modèle comme celui de Reichenbach qui utilise un nombre fini d'attributs et un algorithme unique, on observe une explosion de la complexité.

Mais peut-être que l'approche qui attache des instructions aux marqueurs linguistiques, si elle est critiquable du point de vue de la parcimonie, reste cognitivement plausible. Notre argument principal est celui de l'apprentissage : Des enfants peuvent-ils apprendre ces algorithmes avec la pratique (voir chapitre 2.2 La plausibilité cognitive) ? Ne disposant malheureusement pas de suffisamment de connaissance à propos des capacités d'un cerveau humain sur ce point, nous ne pouvons pas y apporter une réponse définitive.

Il se trouve que le modèle de Gosselin dispose d'un mécanisme supplémentaire : Il arrive que l'exécution des instructions conduise (assez fréquemment) à des **conflits**. Les éléments de représentation ainsi obtenus sont incompatibles entre eux. Par exemple :

(98) *Il découvre le trésor en trois semaines* (procès ponctuel et circonstanciel non ponctuel).

(99) *J'ai terminé dans un quart d'heure* (temps du passé et circonstanciel à valeur de futur).

Interviennent alors des instructions de **résolution des conflits**. Celles-ci vont déformer les représentations (le moins possible) afin que toutes les exigences soient satisfaites. Il existe trois zones déformables :

a) L'intervalle du procès et par son intermédiaire l'intervalle de référence.

b) L'intervalle de l'énonciation.

c) La relation entre l'intervalle de référence et son antécédent.

Ces zones sont déformables à des degrés divers avec les principes de déformation suivants (et un exemple associé à chaque fois) :

a) Déplacement :

(100) *Il est en train de trouver la solution.*

L'intervalle *être en train de* se déplace vers le passé et précède l'intervalle de *trouver la solution*.

qu'on ne puisse exprimer qu'un nombre limité de positions : Pour ordonner des moments, Reichenbach n'a besoin que de deux relations. Les enrichissements de Gosselin introduisent une valeur quantitative pour l'antériorité et des intervalles au lieu de moments, mais on a toujours un nombre fini de résultats.

b) Contraction :

(101) *Il rentra à 11h45.*

L'intervalle de *rentrer* se contracte sur la borne finale, c'est-à-dire pour un intervalle $[i,j]$ on obtient $i \mu j$.

c) Dilatation :

(102) *Il joua du piano pendant dix ans.*

Les bornes de l'intervalle de *jouer du piano* s'écartent pour être simultanées avec celles de *dix ans*, puis est suivie d'une duplication réitérée du procès (Si elle n'a pas lieu on obtient un effet de sens de type habitude).

d) Duplication :

(58) *L'an dernier, Marie mangeait à la cantine.*

L'intervalle de *Marie mangeait à la cantine* est dupliquée un nombre indéterminé de fois.

Contrairement aux instructions des éléments du lexique, les mécanismes de résolution de conflit ne sont attachés "à rien", et sont exécutés (ou s'exécutent) quand c'est nécessaire.

Ces résolutions de conflit agissent sur les éléments de représentation, on peut donc dire que ces derniers sont "dynamiques". C'est à notre avis un point fort du modèle, étant donné que des interprétations d'énoncés peuvent toujours être trouvées, y compris pour les plus farfelues.

Le modèle de Gosselin, bien que l'on ait critiqué son manque de minimalisme (Munch & Dessalles, 2012), reste un modèle complet sur le temps et l'aspect de la langue française. Il traite de nombreux cas et profite des études précédentes pour proposer un modèle cohérent et soucieux d'une approche cognitive. Mais est-il possible de simplifier le modèle ? C'est-à-dire de traiter les mêmes énoncés tout en ayant un modèle plus minimaliste ? **Et surtout, est-il possible de traiter de l'aspect en évitant de placer des instructions dans les éléments du lexique ?**

5.3 Kamp et Reyle

Kamp et Reyle proposent un cadre pour le traitement du sens dans le langage : la Discourse Representation Theory (Kamp & Reyle, 1993), plus communément abrégé en **DRT**. Ils ne se limitent pas au temps et leur système peut être utilisé de manière général pour tout traitement sémantique. Nous allons très rapidement expliquer en quoi consiste la DRT, avant de nous concentrer sur son traitement du temps (temporalité et aspect). Cette seconde partie s'inspire de l'analyse qu'en fait Øyvind Strand (Strand, 2012).

La DRT utilise un système de structures, les Discourse Representation Structures (ou DRS) pour formaliser les représentations mentales du lecteur ou auditeur. Les DRS sont composées de :

- Discourse Referents, qui représentent des entités.
- Conditions, qui représentent des informations à propos des discourse referents.

Prenons par exemple (tiré du premier volume de Kamp & Reyle, 1993, p. 60),

(103) *Jones owns Ulysses. It fascinates him.*

La première phrase peut être représentée de cette manière :

x y
<i>Jones(x)</i> <i>Ulysses(y)</i> <i>x owns y</i>

Figure 13 : DRS de "*Jones owns Ulysses*"

Cette représentation signifie qu'il existe deux référents au discours, x et y ; et trois conditions au discours, Jones, Ulysses et owns ; tel que Jones définit x , Ulysses définit y , et owns définit x et y .

Le système de référents permet de facilement traiter des pronoms référentiels en créant une correspondance entre référents. Si l'on veut obtenir une structure complète de l'exemple (103), on obtient :

x	y	u	v
<i>Jones(x)</i>			
<i>Ulysses(y)</i>			
<i>x owns y</i>			
<i>u = y</i>			
<i>v = x</i>			
<i>u fascinates v</i>			

Figure 14 : DRS de l'exemple (103) complet.

Ici *u* et *v* correspondent aux pronoms *it* et *him*. Ceux-ci se comportent comme des anaphores des référents déjà existants, c'est-à-dire *y* et *x*.

Il se trouve que la DRT utilise le même principe anaphorique pour les analyses temporelles : Le *temps*²⁹ de *fascinates* dans (103) va être rattaché au *temps* de *owns*. La nature de la relation anaphorique dépendra alors de divers éléments, par exemple si la situation est perçue comme un event ou un state.

(104) *A man walked in. He ordered a beer.*

(105) *A man walked in. He was thirsty.*

Strand (2012) illustre cette situation avec les exemples (104), où les deux situations (*to walk* et *to order a beer*) se suivent, et (105), où les deux situations (*to walk* et *to be thirsty*) se chevauchent.

La DRT catégorise alors chaque *temps* de ces situations comme étant un événement (event) ou un état (state). Les *temps* traités seront alors ensuite reliés à un temps de localisation (location time) qui proviendra généralement d'un élément précédent dans le discours. Ce location time correspond généralement à la notion de *temps* de référence introduite précédemment (voir page 58 et page 67), bien qu'il soit possible d'avoir besoin davantage de *temps* que celui de référence pour des phrases plus complexes (les *temps* introduits par des circonstanciels par exemple). Enfin, le *temps* d'énonciation (time of utterance, que l'on notera plus bas "n") est également pris en compte pour lier le discours au contexte courant.

²⁹ Le *temps* que nous mettons ici en italique a ici une valeur bien précise et est considérée comme une entité au même titre que Jones ou Ulysses.

Nous avons vu que l'aspect (event ou state) d'une situation a des conséquences sur les relations entre la situation et le temps de localisation. Dans la DRT, les events (e) seront considérés comme inclus dans leur temps de localisation ($e \subseteq t$), tandis que les states (s) incluront leur temps de localisation ($t \subseteq s$).

La première phrase des exemples (104) et (105) pourront alors être représentée par cette DRS (le symbole "<" signifiant la précédence) :

n	m	e_1	t_1
<p>"a man"(m) $t_1 < n$ $e_1 \subseteq t_1$ $e_1 : walk-in(m)$</p>			

Figure 15 : DRS de "A man walked in".

Nous pouvons noter que la temporalité passé de *walk-in* est représenté en faisant précéder le temps d'énonciation (n) par le temps de localisation (t_1), tandis que l'événement (e_1) lui-même est directement inclus dans le temps de localisation (t_1).

Voyons maintenant comment représenter la seconde phrase de (104) et comment la lier avec la première représentation. Le problème consiste à identifier le temps de référence (r) de la seconde phrase qui doit correspondre à un temps (pour le moment inconnu) de la première représentation. Ce temps de référence sera relié au temps de l'événement (t_2), sans que l'on sache pour le moment de quelle manière. Voici les DRS avant résolution :

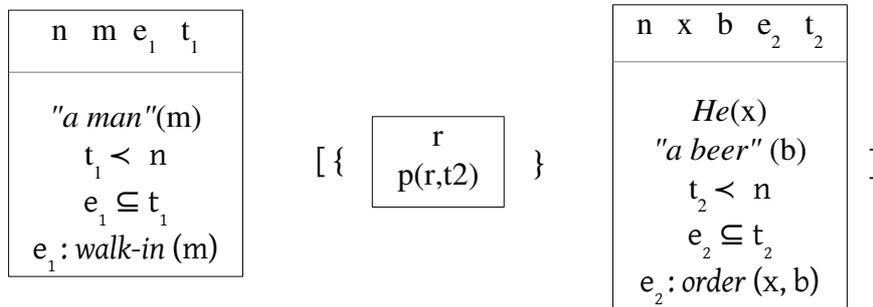


Figure 16 : Unification des DRS de "A man walked in" et "He ordered a beer"

La résolution consiste à identifier r et p (p devant être une relation) par rapport à la première DRS. Dans ce cas, r ne peut correspondre qu'à t₁, qui est le seul *temps* disponible. En ce qui concerne p, la DRT indique qu'attacher un event à un autre event requiert une relation de précédence (<). L'unification (merge) des deux DRS peut alors être réalisée :

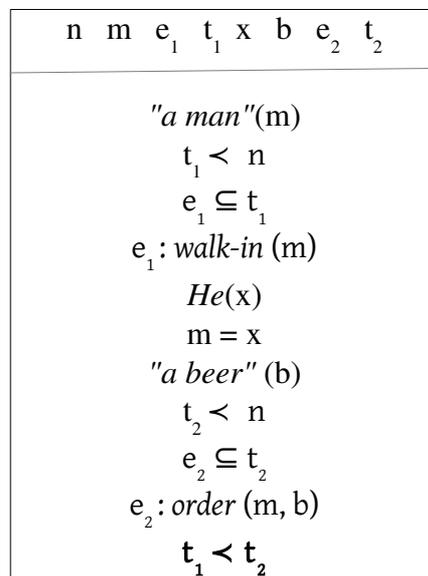


Figure 17 : (104), résultats

Pour le traitement de (105), la seule différence s'appliquera à l'aspect de la seconde situation, qui sera un state plutôt qu'un event, ce qui signifie que ce state (s₂) inclura son temps de localisation (t₂ ⊆ s₂). L'unification des deux DRS conduira alors à avoir le temps de localisation de la première

situation (t_1) incluse dans le temps de localisation de la seconde. ($t_1 \subseteq t_2$ au lieu de $t_1 < t_2$).

Les auteurs de la DRT (Kamp & Reyle, 1993) remarquent néanmoins que le cadre de la DRT ne permet pas de traiter tous les énoncés, par exemple :

(106) *Chris had a fantastic meal. He ate salmon.*

(7) *Marie est tombé. Jean l'a poussé.*

(107) *John turned off the light. The room was pitch dark.*

Dans ces exemples l'organisation temporelle est obtenue grâce à des connaissances/inférences pragmatiques : *Manger du saumon* est une partie du *fantastique repas*, et est donc simultané ou inclus dans ce *repas*, et non pas sa suite. Dans les exemples (7) et (107) les situations décrites par les secondes phrases précèdent ceux des premières phrases, alors qu'un comportement classique (c'est-à-dire sans contre-indications) est de faire suivre temporellement les situations dans l'ordre où elles apparaissent dans le texte.

Il se trouve que la DRT a été enrichie pour justement traiter ce genre de phrase : c'est la Segmented Discourse Representation Theory (SDRT) de Lascarides et Asher (Asher, 1993). Le point important de la SDRT est qu'elle n'utilise pas systématiquement l'aspect (event ou state) pour déduire les relations entre le temps de référence et le temps de éventualité, mais qu'elle utilise aussi les connaissances pragmatiques (qui sont introduites dans la base de connaissance).

La DRT et la SDRT sont des modèles capables d'exprimer beaucoup d'interprétations de temporalité, d'aspect et même de connaissances logiques (sous forme de prédicats) tout en restant relativement parcimonieux. Ils ne lient pas d'algorithmes aux éléments du lexique (comme le font Gosselin (1996) ou de Saussure (2003) pour le français) mais se contentent d'avoir un nombre fixe de règles liées à la syntaxe avec laquelle ils communiquent.

Le principe de liaison sémantique, c'est-à-dire l'obligation de partager des variables entre les prédicats liés syntaxiquement ou discursivement (prédicats au sens logique, comme dans l'exemple (103)), rend le modèle non seulement capable de traiter des éléments de "pragmatique", mais est aussi capable de dépasser le niveau de la phrase : La DRT est capable de traiter des suites d'énoncés.

Ce point fort dissimule néanmoins une faiblesse possible du modèle : celui-ci a alors une croissance monotone de la mémoire et ne précise pas ce qu'il fait des structures construites. La

DRS d'un énoncé en début de paragraphe est-elle toujours intacte au moment où le modèle traite la fin du paragraphe ou d'un livre entier ? Les mémorise-t-il sans mécanisme "d'oubli" au risque d'avoir une explosion des structures ?

Le modèle de Kamp et Reyle semble répondre à nos soucis de minimalisme : Le nombre et la taille des algorithmes semblent maîtrisés, la taille croissante des structures (DRS) n'est pas insurmontables à contrôler, tandis que les entités ontologiques liées à l'aspect (event et state) sont en nombre réduit et fixe.

On peut s'interroger sur l'origine des prédicats (*a beer(X)* ; *order(m, n)* ; etc.). Cette question ne concerne pas directement la partie du traitement de temps et de l'aspect et touche en réalité tout le modèle du traitement du langage : Comment sont structurées les connaissances ?

Si nous n'avons pas de réponse tranchée, nous verrons avec Ghadakpour et Fodor dans le chapitre suivant (6.3.1 Le système conceptuel) que l'utilisation de prédicats n'est pas souhaitable (en tout cas dans la théorie) pour leurs similarité avec un langage of thoughts, critiquable justement pour son manque de parcimonie et de plausibilité cognitive.

Du modèle de Kamp et Reyle nous apprécierons le traitement de l'aspect, conservant notamment l'idée de structure d'informations (les DRS) mais en modifiant quelques attributs clés : Notre structure devra être de taille fixe et chacune des structures instanciées sera temporaire, c'est-à-dire qu'elles pourra être détruite ("oubliée").

5.4 De Swart

Le modèle de l'Aspect de de Swart (1998) est développé dans le cadre plus général de la DRT de Kamp et Reyle (chapitre précédent, p.77). Nous nous basons essentiellement sur la critique qu'en fait Borik (2006) pour en présenter les points essentiels et les défauts.

De Swart propose un modèle que l'on peut représenter comme une structure imbriquée impliquant les trois éléments généralement admis qui véhiculent une information aspectuelle³⁰ :

30 L'astérisque indique que les opérateurs aspectuels peuvent être multiples et récursifs, alors que Tense et eventuality description sont uniques.

[Tense [Aspect* [eventuality description]]]

L'**eventuality description** ("description de la situation") correspond à trois primitives de l'ontologie de de Swart : states, processes et events. On note que ceux-ci se rapprochent des types aspectuels de Vendler, et correspondent à l'aktionsart de la situation.

"There is a straightforward correlation between the aspectual class of an atomic eventuality description and the type of eventuality it denotes: stative sentence introduce states, process sentences refer to processes, and event sentences describe events"

(De Swart, 1998, p. 351)

Elle se sert de l'information sémantique d'homogénéité (que nous présentons p.52) pour séparer states et processes qui sont homogènes, et les events qui sont quantifiables. Elle utilise le progressif pour illustrer ce propos :

(108) *Ann was running* --> Ann Ran

(109) *Ann was running a mile* -\-> Ann ran a mile

Le type de l'éventualité de (108) est un process, c'est un type homogène puisqu'il est possible de passer sans problème du past progressive au simple tense. Le type de l'éventualité de (109) est un event, c'est un type quantifié, où l'inférence précédente n'est plus possible. On entend en effet dans (109) qu'Ann a été interrompue et qu'elle n'a pas pu courir un mile.

Bien que simple et séduisant, ce fonctionnement comporte un problème décrit par Borik : Si une situation est **décrite** comme terminée/bornée, cela ne signifie pas qu'elle l'est effectivement dans le monde.

(110) *From 1994 till 1996 John lived in New York.*

Dans (110), ce n'est pas parce que la situation est présentée comme achevée que *John* ne vit plus à *New York*. De plus, la situation ne peut pas être telique, puisqu'elle réussit les tests d'homogénéité et du progressif :

(111) *From 1994 till 1996 John **was living** in New York. → From 1994 till 1996 John lived in New York.*

(112) *From 1994 till 1996 John lived in New York. → From 1994 till **1995** John lived in New York.*

Le niveau supérieur dans la représentation de de Swart est l'**aspect**. Un nombre d'opérateurs aspectuels peut alors s'appliquer à l'eventuality type, et le transformer en un autre ou le même

type. Par exemple si le type premier de l'eventuality est un event, l'application d'un opérateur progressif en fera un état (state). Le phénomène peut être récursif et plusieurs opérateurs peuvent alors s'appliquer. De Swart donne un exemple avec les opérateurs perfectif et progressif pour l'anglais :

(113) *Jane has been writing a letter*

[PRES [PERF [PROG [Jane write a letter]]]]

Les adverbes temporels sont eux aussi des opérateurs aspectuels, comme *for an hour* et *between 3 and 5*. Ainsi, si la situation *Ann run* est considérée comme un process au premier niveau (eventuality type), la situation décrite par *Ann ran for two hours* est un event.

Borik (2006) critique cette approche en considérant que la situation en elle-même ne va pas changer selon qu'on la décrit d'une manière ou d'une autre. Le système de de Swart ne semble pas capable de faire la différence entre la situation d'origine (du premier niveau) et le résultat final (après l'application du troisième niveau). Cela peut poser problème :

(114) *[Andrew swam a mile] in a week.*

(115) *[Andrew swam for three hours] in a week.*

(116) *(after he saw a cat,) [my dog ran] in a second.*

D'après de Swart, (114) est un event, (115) est event dérivé d'un process après modification de l'adverbe, et en (116) l'eventuality type est un process. Mais selon Borik (La numérotation des exemples est la nôtre) :

If there was no difference between underived and derived events, the predication is that the interpretation of 'in X time' in (114) and (115) should be the same, as opposed to (116).

Interestingly, the example in (115) rather patterns with (116). The only interpretation (116) can have is the following: Andrew had to take part in a swimming competition and he was bad at swimming long distance. As a training program, his instructor told him to practice for three hours every day. The first day, Andrew could swim only for 30 minutes, but after a week of training, he could swim for three hours. In other words, (115) is interpreted as 'It took Andrew a week to start/learn to swim for three hours', Similarly, (116) means 'It took my dog a second to start running (when he saw a cat)'.

Nous renverrons à Borik (2006, p. 113) pour davantage de détails sur cette critique.

Le troisième et dernier niveau de l'information aspectuelle peut être transmise avec le temps, et a

selon de Swart pour application certains temps grammaticaux comme le passé simple et l'imparfait en français. Le premier demande un event quantifiable en entrée, le second n'est acceptable que s'il s'applique à une description homogène et non-quantifiable.

Le modèle de de Swart apparaît comme simple et élégant. Elle évite de lier des suites d'instructions aux éléments du lexique, tout en limitant le champ des solutions possibles en utilisant un nombre définit d'opérateurs. Ceux-ci s'appliquent alors à tous les niveaux de représentation, ce qui permet de n'avoir qu'un seul type d'opérateurs. De plus, la récursivité du système et son caractère imbriqué permettent une très grande expressivité.

Une critique possible du modèle est de "perdre" les informations originales d'aspectualités grammaticales et d'aktionsart en cours de traitement alors qu'elles devraient rester accessibles pour les interprétations finales.

De de Swart on retiendra l'ordonnement de eventuality description, d'aspect et de temps. Le premier élément étant modifié par les opérateurs d'aspect puis de temps. L'ordre de ces imbrications sous-entend une démarche linéaire pour le traitement de l'aspect : C'est-à-dire qu'il semble possible de traiter d'aspect sans obligatoirement prendre le temps en compte. Bien entendu, on s'attendra dans ce cas à des résultats imparfaits selon certains énoncés.

6 Élaboration du modèle

Nous avons présenté une vision générale du traitement du temps dans les langues naturelles et plus particulièrement de l'aspect. D'abord en précisant les deux exigences que nous nous imposons que sont la parcimonie et la plausibilité cognitive , puis en abordant le concept de temps de la logique à la philosophie, en passant par une approche cognitive. Nous avons ensuite introduit les différents concepts et "caractéristiques" du temps dans les langues identifiés par les linguistes et avons présenté quelques modèles du traitement du temps dans le chapitre précédant.

Nous allons dans ce chapitre abordé les travaux non-linguistiques qui ont largement influencé l'élaboration du modèle, notamment ceux de Ghadakpour.

Nous présenterons l'**architecture générale** que nous prêtons au système langagier, détaillerons les liens entre **modules "sémantiques"** et **module "syntaxique"**, puis présenterons le **module "perceptif"**, dont on suppose un fonctionnement très proche du modèle de Gärdenfors (Gärdenfors, 2000, 2014).

6.1 Une approche modulaire

Dans le chapitre 2.2 La plausibilité cognitive nous avons discuté de la plausibilité cognitive et du caractère modulaire supposé de la cognition. Nous avons choisi de suivre cette approche dans la conception de notre modèle, d'abord pour nous rapprocher de cette plausibilité cognitive, ensuite pour ses avantages d'implémentation : La modularité est une caractéristique importante d'un programme informatique. Elle permet de le faire évoluer, d'identifier aisément les erreurs et de favoriser la compréhension du code.

Notre modèle décrit uniquement le module du temps et ses interactions, et pas (ou très peu) les autres modules (en particulier ce que nous appelons le module perceptif). Si nous pouvons nous en contenter pour nos tests en implémentant des versions "simplifiées", la modularité devient impérative si l'on espère un jour passer à l'échelle (traitement de tous les énoncés possibles) en implémentant chacun des modules qui interagissent avec celui du temps. Les théories de la syntaxe et de la perception peuvent évoluer : notre modèle s'y conformera plus facilement.

Chacun des modules que l'on représente dispose d'entrées (inputs) et sorties (outputs) dont le format et la parcimonie sont pris en compte. On cherche en cela à minimiser les communications entre les différents modules par soucis de plausibilité cognitive : Il est plus plausible et rapide de transmettre des informations de quelques bits (du type "oui", "non", "ne sais pas"³¹), que de transmettre des informations plus complexes (listes d'éléments, structures imbriqués, etc.) ou pire encore des algorithmes. **Les échanges d'information entre les modules seront donc minimaux.** Nous supposons possible que chaque module est lié à d'autres, mais nous conterons de définir dans les chapitres suivants les liens entre notre module temporel et les modules lexico-syntaxique et perceptif (pages 90 et 94).

La figure suivante présente les différents modules du modèle dont nous avons besoin pour le traitement du temps. La figure est librement inspirée des travaux de Luc Steels et son équipe sur la Fluid Construction Grammar (Steels, 2011).

31 Ces variables (A, B ou "ne sait pas") seront appelés **switches** dans la suite du document. Voir page 120.

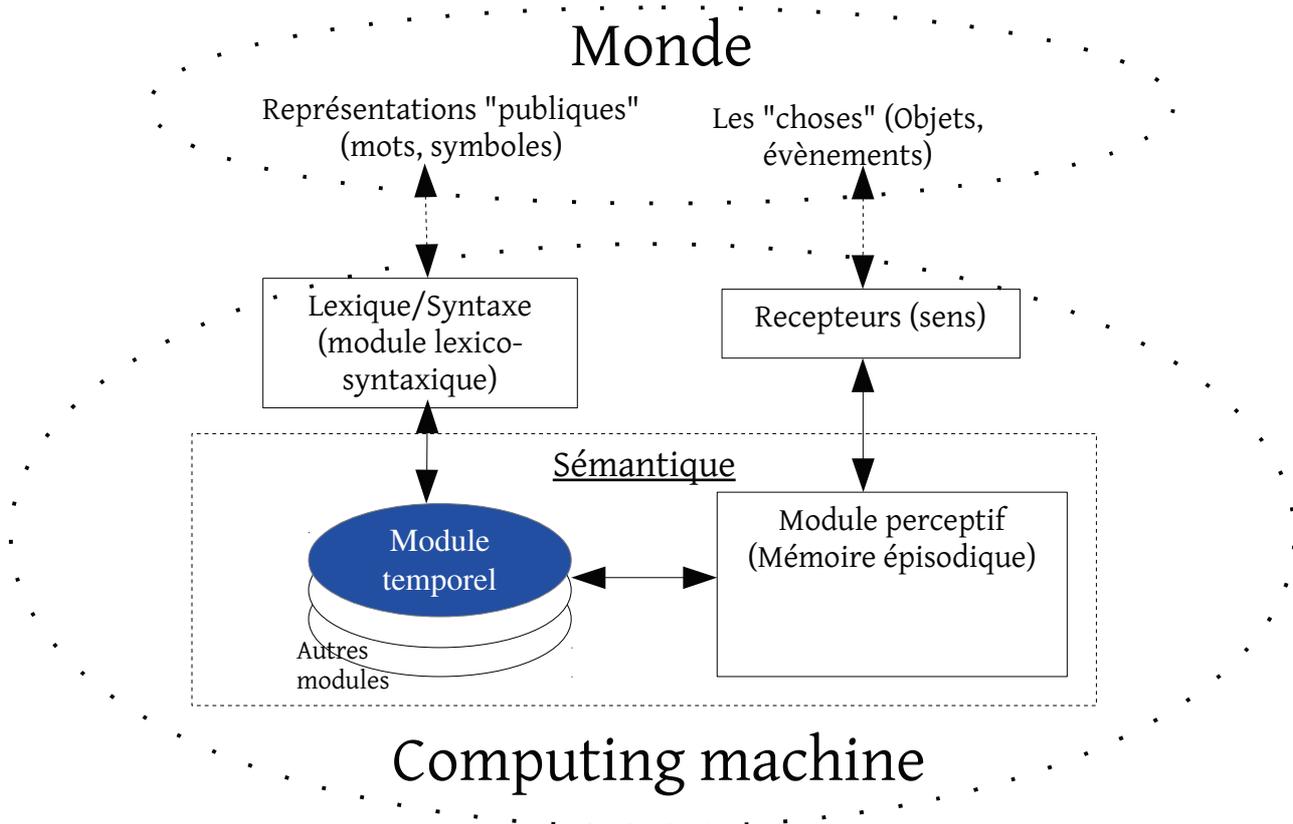


Figure 18 : Une vision du système du traitement du langage

La catégorie "sémantique" ne correspond pas à une entité ou un module et n'est indiquée que pour faire un contraste avec la syntaxe ou la phonologie (qui elle n'apparaît pas dans notre figure).

Les flèches à double sens indiquent les interactions (qui seront minimales) entre les modules. Nous n'avons représenté que les interactions qui nous intéressent (module lexico-syntaxique – module temporel – module perceptif), mais nous considérons que chaque module est potentiellement en relation avec tous les autres.

Nous choisissons de diviser le système cognitif en deux parties : les modules symboliques et les modules sub-symboliques³². Cette distinction va devenir plus claire lorsque nous aborderons le module perceptif un peu plus bas.

32 Un système sub-symbolique est un système où les "atomes" ne sont pas des représentations, au contraire d'un système symbolique où les "atomes" sont des représentations de quelque chose. Dans un système sub-symbolique la signification n'apparaît qu'en considérant un ensemble "d'atomes".

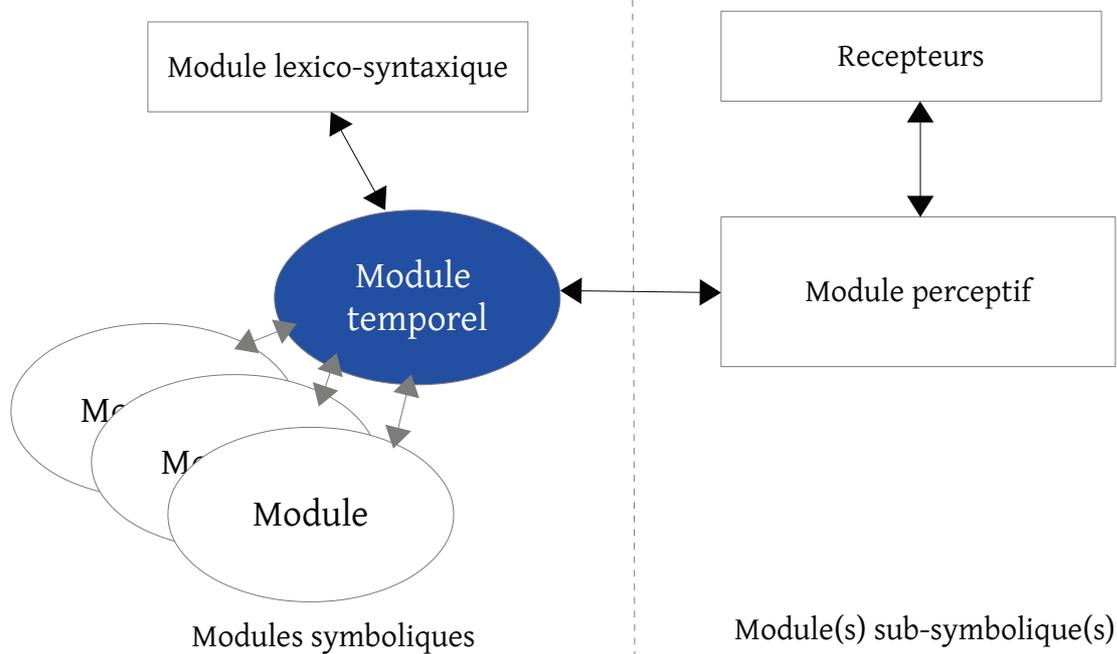


Figure 19 : Division entre modules symboliques et sub-symboliques

Nous distinguons le **module lexico-syntaxique**, qui gère l'identification des lexèmes et leur organisation. La littérature sur le sujet est pléthorique et les implémentations extrêmement variées. Nous n'en proposons pas de nouvelle ici et nous contentons uniquement de lui demander une organisation des unités lexicales, c'est-à-dire un enchaînement de paires lexicales qui formeront un arbre syntaxique. Nous discutons du rôle de ce module dans la partie suivante, La fusion simultanée dans la syntaxe et la sémantique.

La **perception** correspond aux sens (vue, ouïe, etc.) et à leur transformation en une information stockable par la mémoire. Sa sortie prend(ra)it la forme "d'images perceptives", et de liens entre elles. Son fonctionnement nous est largement inconnu et tout comme le module lexico-syntaxique nous ne nous prétendons pas en proposer une implémentation élaborée. Nous nous contenterons d'une version très simplifiée des résultats sous la forme du **module perceptif**³³.

La mémoire perceptive est inspirée de la mémoire épisodique (Tulving, 1983). On suppose que le module la représentant se comporte comme une base de données gérant des images perceptives et les liens entre elles. Ces liens peuvent être de différents types, comme "est", "est une partie de", "précède", etc. Bien entendu, ce sont les liens de nature temporelle et causale qui nous intéressent le plus. Nous détaillons ce module (qui n'est certainement qu'une vague et naïve approximation de la mémoire épisodique humaine !) dans le chapitre 6.3 "Le module perceptif" (page 94).

33 Le module perceptif est décrit comme ayant un fonctionnement sub-symbolique, mais une implémentation plus classique de type symbolique, comme une ontologie, peut être suffisante dans notre contexte. Nous y reviendrons.

Entre le module lexico-syntaxique et le module perceptif se trouve toute une série de modules dit "sémantiques symboliques". Le module perceptif traite des informations de type sub-symboliques (Nilsson, 1998), tandis que les modules sémantiques non perceptifs utilisent des informations symboliques. De notre côté, le module perceptif se comporte comme une boîte noire, approche peu satisfaisante mais que nous sommes contraints de prendre pas soucis de simplicité. Notons que Gärdenfors (2000) a formalisé un modèle conceptuel non-symbolique intégré à la perception, modèle qui pourrait être utilisé pour notre module perceptif (nous en reparlons page 94).

Parmi ces modules sémantiques symboliques, nous supposons l'existence de modules bien définis et parcimonieux : module de gestion de l'espace, de causalité, de négation, de modalité, du temps, etc. Nous appelons **module du temps**, celui responsable du traitement de la temporalité et de l'aspect, bien que nous nous concentrerons par la suite sur l'aspect. Ce module est le cœur de nos travaux, c'est lui qui réalisera les calculs sur les informations temporelles obtenus des modules syntaxiques et perceptif. Nous le détaillons dans le chapitre 7 Description du modèle, page 115.

On remarquera que les flèches reliant les différents modules de notre figure sont bidirectionnelles. Nous voulons avoir le même modèle pour gérer **l'interprétation** – qui prend un énoncé en entrée et fournit des valeurs "sémantiques" (voir chapitre 7 Description du modèle, page 115, pour les détails sur ses valeurs) ; ET **la production** – qui prend des informations sémantiques en entrée et construit un énoncé³⁴. Il nous semble peu plausible et pas du tout minimal d'avoir deux systèmes distincts pour gérer interprétation et production.

6.2 La fusion simultanée dans la syntaxe et la sémantique

Notre sujet n'est pas la syntaxe, mais le traitement du temps. Il est néanmoins obligatoire de passer par la syntaxe pour disposer de quelques informations nécessaires à toute tentative de traitement sémantique. Nous n'allons néanmoins pas détailler le fonctionnement du module syntaxique. Comme nous l'avons déjà dit, de nombreux travaux existent déjà sur le sujet et

³⁴ N'ayant pas implémenté ce circuit, nous ne donnerons pas les informations précises en entrée, sachant que celles-ci ne correspondent pas exactement aux valeurs sémantiques de sortie de l'interprétation.

continuent d'être réalisés.

Comment choisir le ou les modèles qui correspondent à notre vision du traitement du langage ? On voudrait bien évidemment qu'ils suivent nos mesures de parcimonie et plausibilité cognitives, mais nos contraintes ne portent pas tant sur le fonctionnement interne du module, que sur ses entrées et sorties.

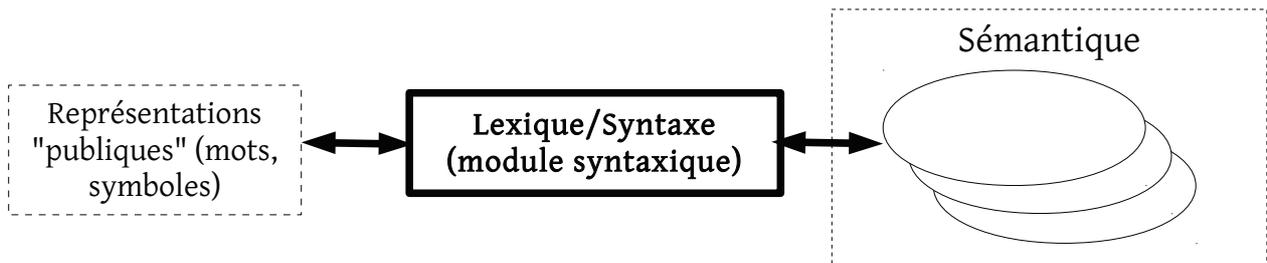


Figure 20 : Interfaces du module syntaxique

Du côté du "Monde", il y a les **énoncés** (écrits et oraux – Dans ce dernier cas il est nécessaire d'intégrer un module phonologique au système, qui n'apparaît pas sur nos figures), que nous comprenons dans un sens très large comme "une unité linguistique constituée d'une suite ordonnée de mots ou symboles". L'énoncé est grammaticalement correct si le module syntaxique est capable d'en tirer une structure syntaxique. Voici par exemple deux arbres syntaxiques d'un énoncé simple :

(117) *Jean frappe la balle*

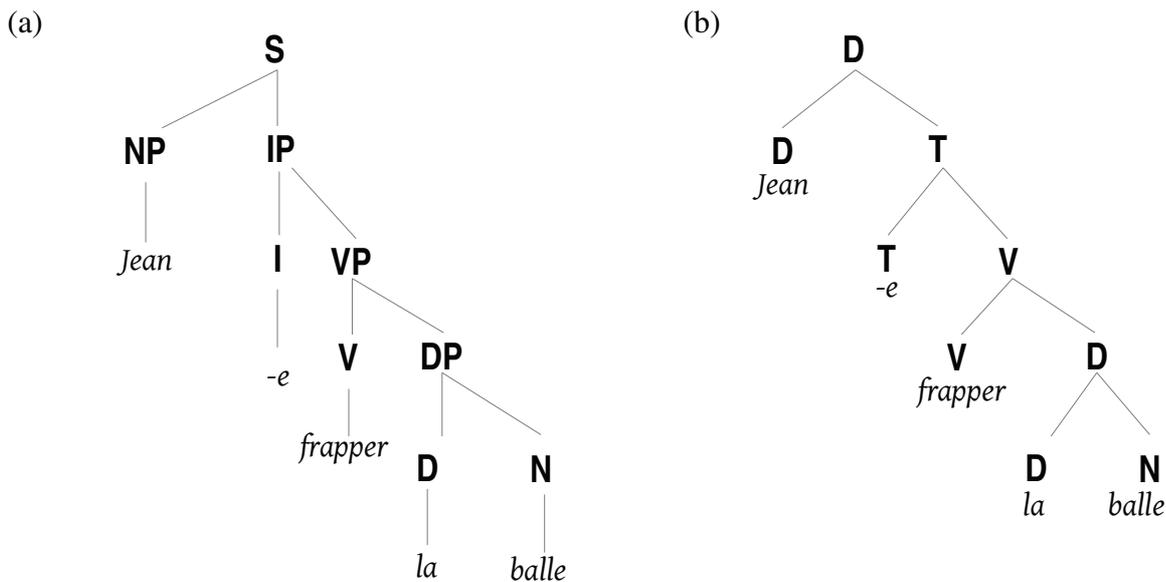


Figure 21 : Représentations syntaxiques de l'exemple (117)

(a) est un arbre syntaxique construit selon la théorie X-bar (Chomsky, 1970), tandis que (b) est construit en suivant la théorie du programme minimaliste (Chomsky, 1995). Ces arbres ne sont que des représentations possibles et les théories syntaxiques encore en évolution pourraient bien en proposer de nouvelles. Nos modules sémantiques n'ont pas besoin de beaucoup d'informations et peuvent s'arranger de (a) et de (b) comme nous allons le voir un peu plus bas.

Les arbres sont ici présentés dans leur totalité, mais la quantité d'informations est assez conséquente et bien loin de notre objectif déclaré d'avoir des transmissions d'informations entre modules minimales. De plus on devrait se demander si le module syntaxique transmet l'information uniquement lorsqu'un énoncé est "complet", c'est-à-dire qu'un S (Sentence) a été trouvé ? Que ce passerait-il si, comme dans les situations courantes, on se retrouve avec des énoncés à rallonge et parfois grammaticalement "incorrects" ? Voici un exemple de discussion :

Contexte : B et C veulent acheter une voiture break. A fait observer que ce genre de véhicule est cher.

A1- *Mais, heu, quand je disais une voiture spacieuse,... Ah oui, break. Évidemment, il y a de la place dans un break. Seulement les breaks sont rares et chers, je crois, ils sont recherchés.*

B1- *Non, non, ou ils sont*

C1- *Non, j'en ai vu une à vendre, mais il [B] a pas voulu acheter.*

A2- *et ensuite, ils ont des kilométrages impressionnants. Ils ont 200.000 bornes quand tu les...*

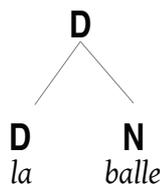
Dans la pratique, on se rend bien compte que l'on comprend les mots avant d'avoir un énoncé complet et bien formé.

Notre hypothèse sur les communications entre le module syntaxique et les modules sémantiques est la suivante :

Le module syntaxique donne un couple d'unités lexicales ou syntaxiques ordonnées (l'un des membres domine) à la sémantique dès que le couple a été identifié.

Dans le cas de l'exemple (117), le premier couple syntaxique complet repéré est "la" et "balle". Qu'importe que l'ensemble soit un 'DP' ou un 'D', composé d'un déterminant et d'un nom : l'information est purement syntaxique et ne sera pas transmis. Les syntacticiens (Chomsky, 1995) parlent de fusion³⁵ syntaxique (syntactic merge) lorsque deux unités lexicales ou syntaxiques sont associées. Dans notre exemple (b), 'D' et 'N' sont fusionnés en un 'D'. Cette fusion syntaxique, suite à notre hypothèse, implique alors une réception simultanée par les modules sémantiques. Nous considérerons ici que le traitement réalisé par les modules sémantiques sera réalisé immédiatement. Il se trouve que (dans notre modèle) la sémantique réalise elle aussi un mécanisme de fusion, que l'on appellera fusion sémantique (semantic merge). Le détail de cette fusion sémantique sera détaillé dans le chapitre 7.3 "aMerge : La fusion sémantique (aspectuelle)" (page 129).

Ce qu'identifie le module syntaxique :



Ce qui est transmis aux modules sémantiques :

La (dominant), **balle** (complément)

Figure 22 : Sortie du module syntaxique

Nous n'affirmons pas ici que les données échangées entre module syntaxique et modules sémantiques se limite à une paire d'éléments hiérarchisés. Notre modèle du temps n'a pas besoin de plus, mais il est probable que d'autres éléments soient nécessaires dans un modèle plus global du traitement du langage, particulièrement lorsque l'on souhaite produire des énoncés en partant du sens.

35 Nous avons choisi de traduire le to merge anglais par fusionner plutôt qu'unifier pour nous démarquer de l'unification logique.

Nous supposons une corrélation directe entre syntaxe et sémantique, mais la syntaxe n'est pas forcément la seule organisatrice du traitement sémantique. Certains phénomènes comme l'amorçage ou les structures discursives seraient traitées en considérant les capacités de stockage des différents modules sémantiques : Il est envisageable que ces modules (y compris le module perceptif) puissent garder des informations en mémoire et ne les réutiliser que plus tard.

Nous ne traitons pas de ces phénomènes dans le document, mais considérons que l'interface directe entre syntaxe et sémantique ne les empêche pas.

6.3 Le module perceptif

Endel Tulving (1983) a développé le concept de mémoire épisodique. Basiquement, la mémoire épisodique inclut la mémoire des événements autobiographiques, en contraste avec la mémoire procédurale et la mémoire encyclopédique. Tulving fait la distinction entre savoir (knowing), qu'il lie à la mémoire encyclopédique, et se rappeler (remembering), qu'il lie à la mémoire épisodique. Tulving a identifié trois propriétés primordiales d'un souvenir de la mémoire épisodique : une perception subjective du temps (subjective sense of time, c'est-à-dire la capacité de voyager mentalement à travers le temps), la connexion au soi (connection to the self), et la conscience auto-noétique (C'est-à-dire la conviction d'avoir soi-même vécu l'événement). Ces souvenirs comprennent les images visuelles (visual imagery), la structure narrative, l'identification d'informations sémantiques et enfin la sensation de familiarité (Hassabis & Maguire, 2007).

Nous mettons de côté la sensation de familiarité pour nous concentrer sur les autres attributs des souvenirs : Nous associons à la structure narrative la capacité d'ordonnement (chapitre 6.3.3), et les notions d'images visuelles et d'identification d'informations sémantiques à l'image perceptive (chapitre 6.3.2).

6.3.1 Le système conceptuel

Avant de détailler ce que nous attendons du module perceptif, il nous semble nécessaire de

clarifier la notion de "concept". Il y a au moins deux définitions des concepts :

Dans la tradition empirique, où le concept est une idée ou une représentation de l'esprit qui résume plusieurs objets empiriques ou mentaux (par abstraction ou généralisation de traits communs par exemple).

Et dans une version physicaliste en philosophie du langage, où le concept est une "représentation mentale" et correspond à une entité dans le cerveau. Une entité serait en pratique un symbole ou groupe de symboles à caractère permanent. Dans un système de prédicats par exemple, le concept d'un objet serait l'ensemble des prédicats qui lui sont attachés. Le concept de chien serait alors représenté par le **symbole** DOG (et son sens correspondrait à la liste des prédicats attachés : "a 4 pattes", "a de la fourrure", "est un mammifère", etc.).

C'est la seconde définition, celle des concepts comme représentations mentales, que nous allons discuter dans la suite.

Comme nous l'avons vu précédemment (page 86), nous avons choisi de considérer le module perceptif comme un module sub-symbolique, et non pas symbolique comme c'est le cas pour une ontologie ou une liste de prédicats. Dans la pratique (dans notre implémentation), nous garderons un module perceptif symbolique par soucis de simplicité, tout en gardant à l'esprit que celui-ci "devrait" être sub-symbolique.

Mais pourquoi ce choix ? Pourquoi ce rejet du symbolisme (Ghadakpour parle aussi de "système moléculaire") du système conceptuel ?

Nous présentons quelques-unes des critiques de Ghadakpour (2003), qui elle-même s'inspire directement de Fodor et al. (1980), pour remettre en cause la validité des systèmes conceptuels symboliques :

- Le fait de doter les concepts d'une structure interne, par exemple de type ontologique, est motivé par le souci d'assurer la compositionnalité. La compositionnalité permet de calculer le sens d'un composé à partir du sens des mots qui le constituent. Or, la compositionnalité structurelle est monotone. Cela signifie que la composition du sens prévoit des structures qui croissent en taille et en complexité à mesure que le discours évolue. Or les mécanismes de fusion structurelle n'incluent aucun dispositif "d'oubli", si bien que les structures conceptuelles risquent de croître indéfiniment au fur et à mesure des compositions.

- Les structures lexicales sémantiques doivent être ancrées dans la perception. Cela implique que

les structures perceptuelles doivent être répliquées comme structures conceptuelles. Une telle répllication est nécessairement partielle et imprécise, alors que nous avons une connaissance **omnipotente** du sens. Nous sommes capables de penser sur toutes les subtilités que nous fournit notre perception ; toute perception peut en principe conduire à une conceptualisation. Au contraire, les structures lexicales conceptuelles sont nécessairement finies en variété et en précision.

- Cette façon d'expliquer la composition du sens revient à réaliser une traduction d'expressions linguistiques vers des structures conceptuelles. Ce langage "mental" demande un lexique, une ontologie conceptuelle par exemple. Pourtant les tentatives pour concevoir une telle ontologie rencontre des problèmes considérables, tandis que son comportement systématique n'est pas expliqué (Fodor, 1998).

L'alternative consiste à considérer que la représentation du sens est perceptive. Gärdenfors (Gärdenfors, 2000, 2014) propose que les significations (meanings) sont des constructions géométriques. Ceux-ci appartiennent à des espaces métriques qu'ils partagent avec les perceptions. En conséquence, deux significations du même espace conceptuel sont plus ou moins similaires. Plus important, la signification lexicale se rapporte à une région convexe de ses espaces conceptuels. Cela signifie que si deux objets sont appelés "livre", tous les objets situés dans leur intervalle dans l'espace conceptuel seront également identifiés comme des "livres". Gärdenfors cherche à identifier les différents espaces dans lesquels les significations lexicales sont situés, en fonction de leur nature sémantique (comme les événements, actions, qualités) ou de leur rôle syntaxique (noms, adjectives, verbes). Il observe par exemple que les noms correspondent à des régions dans des espaces multi-dimensionnels, alors que les adjectifs font plutôt référence à des régions avec peu de dimensions, voire même une seule dimension.

Le modèle de Gärdenfors permet de réaliser des catégorisations, c'est-à-dire distinguer si un objet appartient ou pas à une région conceptuel, en se justifiant sur son appartenance à l'un ou l'autre des espaces.

Ghadakpour, qui partage cette approche, se demande comment résoudre le problème de l'interface entre ce système conceptuel perceptif et les phénomènes de raisonnement et de logique que l'on observe dans le langage : elle propose (2003) l'utilisation d'un **opérateur de contraste** qui permettrait d'appliquer des relations topologiques sur des représentations

perceptuelles à plusieurs dimensions, et de produire des représentations symboliques transitoires utilisables par un appareil "logique". En effet, si l'on se restreint aux espaces conceptuel de Gärdenfors il serait a priori possible de faire appartenir à une région conceptuelle (une catégorie) n'importe quel objet, pourvu que l'on soit suffisamment flexible sur les distances entre objets prototypiques et objets comparés : Tous les objets sont alors plus ou moins des livres.

Un opérateur de contraste qui permettrait, en comparant un objet à son plus proche prototype, d'obtenir un vecteur à peu de dimensions. D'après Gärdenfors ce vecteur à peu de dimensions tombera certainement dans la région des adjectifs. Ainsi en comparant un livre de 20 pages à un livre prototypique apparaîtrait uniquement la dimension d'épaisseur.

Dans notre modèle, le module perceptif correspond au système conceptuel de Ghadakpour, avec la capacité de s'interfacer avec un système symbolique. Ce "système symbolique" correspond chez nous aux "modules sémantiques symboliques", tel que celui du temps (voir Figure 19, page 89).

En pratique nous sommes tout aussi contraints que les autres modèles sur le traitement du langage : il n'existe pas encore (à notre connaissance) d'implémentations d'un système conceptuel sub-symbolique garantissant la systématisme et la compositionnalité ; Nous n'en proposerons pas un ici.

Notre sujet restant le temps, nous nous concentrerons sur le module temporel et ses interactions. Nous décrirons dans les chapitres qui suivent les attributs minimaux que le module perceptif devra avoir afin de pouvoir interagir avec le module temporel. L'implémentation du module perceptif ne doit être vue que comme une simplification extrême des résultats que pourraient fournir un système plus complet (par exemple celui de Gärdenfors). Mais il serait certainement envisageable d'utiliser un système avec des prédicats ou des ontologies comme solution temporaire, en "trichant" sur le fonctionnement de l'image perceptive.

6.3.2 L'image perceptive

L'image perceptive, ou simplement **image** dans la suite du document, est un élément du module perceptif (voir Figure 18 : Une vision du système du traitement du langage). Elle fait directement référence à la théorie de Stephen M. Kosslyn (1996) mais ne prétend absolument pas en être une implémentation : En effet l'imagerie mentale reste un problème non résolu et qui sort de notre

domaine de compétence, aussi nous nous contenterons ici d'en présenter une simplification nécessaire pour le traitement du temps et de l'aspect.

Pour se faire une idée de ce que nous entendons par image perceptive, imaginons une scène, une plage par exemple. Vous pouvez mentalement visualiser cette plage et la compléter en y rajoutant des détails : des baigneurs, le soleil, des palmiers du sable fin, etc. Nous sommes capables de nous "rapprocher" d'un élément par l'imagination. La scène elle-même est une image, mais chacun des éléments la composant est aussi représenté par une image qui peut être indépendante des autres. Nous ne savons pas comment le module perceptif lie les images entre elles, ni comment il est capable de les construire ou déconstruire, mais cela nous suffira pour le moment.

Dans notre implémentation, une image perceptive est représentée par un symbole, un atom prolog. C'est une simplification qui va nous permettre de proposer un module perceptif capable d'interagir avec le module temporel. On se contente de reprendre l'idée de "réalisme" des concepts (que nous critiquions précédemment...): le symbole (l'atom) "repas" est l'image perceptive d'un repas prototypique.

6.3.2.1 La fusion perceptive

L'une des particularités des images qui nous intéresse est leur capacité à être **fusionnées** avec une importante flexibilité. Associez les mots *repas* et *cantine*, et vous obtiendrez ("vous vous imaginerez") une image radicalement différente de l'association de *repas* et *plage*, ou de *repas* et *chien*.

Nous considérons que le module perceptif fonctionne de la même manière : Il prend deux images et en produit une nouvelle.

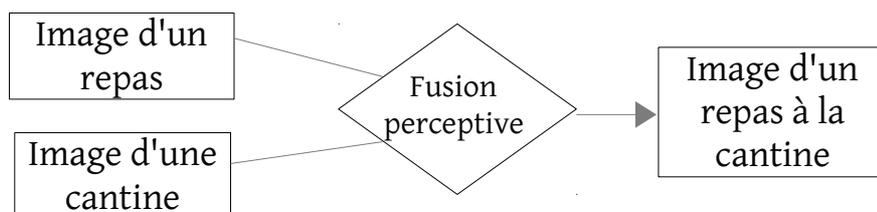


Figure 23 : Fusion perceptive

On constate avec intérêt que le schéma : "deux éléments en entrée" - "une fusion" - "un élément en sortie" est le même que pour la syntaxe. De plus nous utilisons le terme de fusion (merge) dans les deux cas. Nous considérons que :

La fusion qui s'opère au niveau syntaxique guide les fusions des modules sémantiques et perceptifs. Ces fusions sont faites en parallèle.

Voir 7.3.2.1 Unifications avec contraintes perceptives, page 133, pour les détails de la fusion perceptive.

6.3.2.2 Images et informations sémantiques

Nous considérerons que le terme de "sémantique" inclut les informations sémantiques du module perceptif ET les informations symboliques non-perceptives.

Mais que se passe-t-il lorsque la syntaxe fusionne *la* et *plage*? Est-ce que *la* dispose individuellement d'une image au niveau perceptif? C'est très improbable, pourtant il est indéniable que *la* transporte une information sémantique : Détermination, genre, etc.

La fusion perceptive a lieu en même temps que la fusion syntaxique : Nous supposons que le module perceptif fusionnera l'image de *plage* avec une absence d'image. On obtiendra alors une nouvelle image de *plage*. Nous insistons sur le fait qu'il s'agit d'une **nouvelle** image : Il est fort probable que nous disposions d'une image d'une plage prototypique, alors que les informations sémantiques apportées par *la*, notamment la détermination, ont pu amener le module perceptif à renvoyer une ou plusieurs images différentes.

Nous venons de supposer que *la* apporte des informations sémantiques (la détermination) qui ne sont pas des images perceptives. Existe-t-il des mots et lemmes qui apportent des **informations sémantiques temporelles**? A priori oui : ce serait le cas des temps de conjugaisons ou des prépositions comme *pendant* ou *en*³⁶. Nous verrons en détails ces informations temporelles dans le chapitre AIS : La Structure d'Information Aspectuelle(page 117).

Pour notre exemple, *la* disposerait d'informations sémantiques symboliques, stockées dans une

³⁶ Nous discutons davantage des images possibles de *pendant* et *en* un peu plus bas, dans le chapitre 6.3.3.2 Des images relationnelles (page 109).

structure appropriée (voir chapitre 7.2 AIS : La Structure d'Information Aspectuelle, page 117). Mais en l'absence d'image perceptive, *la* n'aurait pas d'informations sémantiques sub-symboliques.

En sortant ces informations sémantiques du module perceptif nous soulevons un problème : Que faut-il garder dans la perception, et que doit-on extraire qui puisse être considéré comme "informations sémantiques" ? Si l'image est dans la perception, est-ce que les qualités d'un objet qu'il représente se trouvent dans la perception ou en-dehors ? Par exemple, pour l'image de *balle rouge*", comment traiter la propriété "rouge" ? De "rondeur" ? Plus proche de notre sujet, où placer les informations dites temporelles comme la durée ou les valeurs aspectuelles ?

Nous proposons de considérer comme informations sémantiques les attributs qui peuvent être obtenus ou calculés à partir d'un mot ou lemme donné (sans faire appel au contexte et aux connaissances du monde), et comme informations perceptives tout le reste.

Par exemple la durée d'un repas dépend des connaissances ou du contexte et n'est pas du tout fournie par le seul mot *repas*. En revanche la conjugaison du futur simple en français apporte toujours la même valeur de temporalité, qui est celle du futur. Attention cependant à bien distinguer l'information fournie par le lexique, est celle obtenue comme résultat final.

En effet informations en entrée et en sortie ne sont pas les mêmes si l'on considère que les interprétations sont obtenues après un calcul (et pas seulement par combinaisons de structures).

(118) *Marie arrive en train.*

Dans l'exemple (118) *Marie* peut être en train d'arriver, c'est-à-dire que l'on observe le moment où le train s'arrête en gare, elle peut être encore dans le train en mouvement une heure avant son entrée en gare, elle peut même encore être encore dans son lit et n'arriver à destination que le lendemain.

Comment distinguer les valeurs obtenues avec le mot/lemme, celles qui sont fournies par la perception, et celles qui sont le résultat d'opérations ? Impossible à dire à ce stade. Nous défendrons dans la suite les choix que nous avons faits en situant dans la "sémantique" ou la "perception", tel ou tel attribut. Les informations perceptuelles sont la durée (chapitre 6.3.2.3) et la "classe" aspectuelle (chapitre 6.3.2.4). Les informations sémantiques sont présentées dans le chapitre 7.2 AIS : La Structure d'Information Aspectuelle

6.3.2.3 La durée

Nous avons déjà introduit les notions de durée et d'ordre de grandeur dans le chapitre [4.1.3 Ordre de grandeur](#), mais allons maintenant préciser notre approche et les choix faits pour l'implémentation de ce phénomène.

L'être humain est capable d'estimer la durée d'un événement. Sa valeur est rarement très précise et est sujette au contexte. Pour reprendre l'exemple du *repas*, celui-ci pourra durer plusieurs heures dans le cas d'un banquet de mariage, ou quelques minutes s'il n'est constitué que d'un simple sandwich. Ces valeurs pouvant atteindre des extrêmes si on parle du repas d'une fourmi, ou de celui de Pantagruel. Dans de nombreux cas il n'existe pas de durée du tout : l'image de *cercle* n'en a pas³⁷.

Pourtant cette durée n'est pas libre pour autant et il existe une contrainte qui rendra "bizarres" certains énoncés.

(119) *Elle a mangé son plat en dix minutes.*

(120) ?? *Elle a mangé son plat en deux ans.*

(121) ?? *Il a fait du vélo deux minutes après son voyage en Chine.*

Si (119) est parfaitement acceptable, (120) l'est beaucoup moins. Il n'y a aucun problème de syntaxe ici, mais un conflit entre la durée envisageable pour manger un plat et les deux ans qu'on lui prête³⁸.

Les trois exemples précédents font tous intervenir une durée explicite (*dix minutes, deux ans*) qu'il ne semble pas possible d'adapter à la durée prêtée à la situation de *manger son plat* ou d'un *voyage en Chine*. Y a-t-il une différence de traitement au niveau perceptif entre les deux cas ? Peut-être que *manger son plat* dispose d'une image, et que *dix minutes* n'en a pas ? Comment est représentée une minute, ou une heure ? On pourrait préciser à moitié en plaisantant que *dix minutes* n'a pas la même longueur pour tout le monde. Cela signifie-t-il que *minute* dispose aussi d'une image perceptive, mais que celle-ci serait moins flexible (à cause de la standardisation des horloges) ? La possibilité de comprendre des énoncés comme *dans trois tours de pistes* ou *en six tics d'horloges* nous

³⁷ Voir chapitre 7.2.6.

³⁸ La phrase redevient acceptable si on situe la scène dans un contexte mythologique, si "elle" est un dragon ou une déesse.

le laisse fortement penser.

Attention néanmoins, ce serait une erreur de ne considérer que la dimension "durée" d'une situation pour en déterminer la comptabilité.

(122) *Elle a mangé du chocolat pendant dix minutes.*

(123) *Elle a mangé du chocolat pendant ces dix minutes.*

(124) *Elle a mangé du chocolat pendant deux ans.*

Ces trois exemples sont très différents de (119) et (120) par l'action (*manger du chocolat* au lieu de *manger son repas*), par la préposition (*pendant* au lieu de *en*) et par l'utilisation dans (123) du déterminant *ces*. Un phénomène de répétition apparaît pour (124) ; dans (123) l'action peut être simultanée avec la totalité des dix minutes **ou une partie** seulement, chose qui n'est pas possible dans (122).

Nous n'allons pas rentrer immédiatement dans l'explication de ces phénomènes, ce qui sera fait dans le chapitre 7.3.3 Opérations, mais allons tirer quelques conclusions sur un traitement possible de la durée par le modèle :

- 1) Les durées de nombreuses images sont très flexibles et dépendent du contexte.
- 2) La durée d'une image semble pouvoir être donnée à partir de la quantité de la durée d'une autre image : Une heure est égale à soixante minutes.
- 3) Le modèle doit pouvoir comparer deux images et donner une réponse positive ou négative sur la compatibilité de leurs durées.
- 4) La comparaison des durées de deux images semble proche de la comparaison du logarithme de leur durée (dans la même "unité") prototypique. On dira que l'on compare leurs **ordres de grandeur**.

Selon cette proposition, pour calculer l'ordre de grandeur d'une image, il est nécessaire de disposer de sa durée prototypique. Nous avons choisi de ne pas implémenter celle-ci mais de fournir directement l'ordre de grandeur. Celui-ci sera fixé de manière un peu arbitraire, mais ce sera largement suffisant pour remplir le point 3) précédent.

Les points 1 et 2 seront largement simplifiés lors de l'implémentation en attribuant une valeur en secondes à l'image. Ceci est uniquement un choix d'implémentation et nous sommes bien conscient que le module perceptif doit disposer d'un mécanisme d'évaluation de l'ordre de grandeur très différent.

Par exemple l'image de *repas* aura un ordre de grandeur de 3,5. C'est-à-dire une approximation du logarithme en base 10 de 3 600 (secondes, soit 1 heure, valeur que l'on a choisi de lui donner). Le choix de la seconde pour réaliser le calcul de l'ordre de grandeur n'a pas d'importance, on aurait aussi bien pu choisir la minute : En effet le résultat final est une comparaison entre deux ordres de grandeurs calculés à partir de la même unité. Voici quelques autres exemples d'ordre de grandeur pour des images perceptives :

- *cligner des yeux* : -1
- *croquer dans une pomme* : 0
- *prononcer une phrase* : 0,7
- *courir cent mètres* : 1,1
- *manger un repas* : 3,5
- *un hiver* : 6,9
- *âge de l'univers* : 17,6

Déterminer à partir de quelle différence les ordres de grandeur sont incompatibles est une gageure. Nous avons fixé cette valeur à 1. Ainsi on considère difficile de faire correspondre *croquer une pomme* et *courir cent mètres*. Par contre il est concevable de *prononcer une phrase* et *courir cent mètres* simultanément (de faire correspondre leur début et fin).

Voici une schématisation de la communication entre le module temporel et le module perceptif³⁹ :

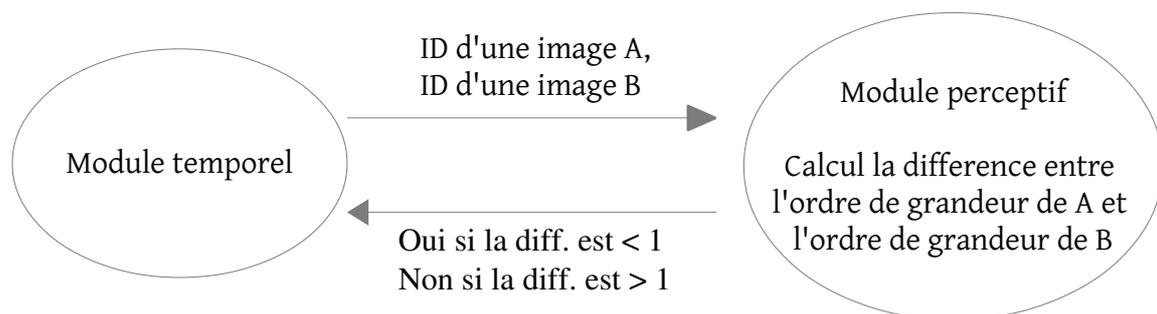


Figure 24 : Communications entre modules temporel et perceptif

³⁹ "ID" correspond à identifiant. Nous expliquons le fonctionnement de cet attribut dans le chapitre 7.2.3 ImageID.

6.3.2.4 Le type aspectuel

Nous avons présenté les classes aspectuelles dans le chapitre [4.1.1 Classes aspectuelles](#) et reproduisons ici la liste de Vendler avec leurs propriétés et des exemples :

- Les **états** (state) sont statiques, duratifs et atéliques

Ex : *be in the house.*

- Les **accomplissements** (accomplishment) sont dynamiques, duratifs et téliques.

Ex : *build a bridge.*

- Les "**achèvements**" (achievement) qui sont dynamiques, instantanés et téliques.

Ex : *realize an error.*

- Les **activités** (activity) qui sont dynamiques, duratifs et atéliques.

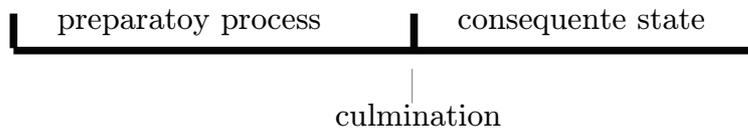
Ex : *walks.*

Dans les chapitres 4.1.2, 4.1.4 et 4.1.6 nous avons discuté des caractéristiques choisies pour identifier les classes aspectuelles, c'est-à-dire respectivement la durativité, la dynamicité et la télicité.

La durativité est une caractéristique différente de la durée (ou plutôt de l'ordre de grandeur) que l'on a décrit précédemment puisqu'elle est une valeur binaire qui indique si l'image a une existence temporelle (et donc un ordre de grandeur) ou pas. L'image du prototype de *pomme* n'a a priori pas d'existence temporelle. L'exemple de Vendler réaliser une erreur (*realize an error*) est plus ambigu : On peut prendre du temps pour réaliser une erreur... Est-ce que cela signifie qu'il existe deux sens à *réaliser une erreur* ? L'un duratif et l'autre instantané ? Il se trouve que pour cet exemple, nous n'avons pas besoin de cette information pour obtenir un calcul correct de l'aspect : "L'instantané" de Vendler est, dans notre modèle, le résultat d'une opération de "prédication" (que nous explicitons dans le chapitre [7.3.3.2 Focus-out et "prédication" temporelle](#), page 138). Cette prédication a pour effet de sortir la situation, ici *réaliser une erreur*, de son sens temporel pour la placer sur une autre dimension sémantique (*réaliser une erreur* fait écho à *NE PAS réaliser une erreur*).

Moens et Steedman (1988) montrent que l'interprétation d'une situation comme étant une

activité ou un achèvement (ils parlent de culmination) est dépendante d'une référence temporelle, qui indique à quelle partie ("nucleus") on fait référence.



Deux représentations possibles selon Moens et Steedman de l'exemple (91) :

- (91) *When they built the 39th Street bridge ...*
- a. *... a local architect drew up the plans.*
 - b. *... they used the best materials.*
 - c. *... they solved most of their traffic problems.*

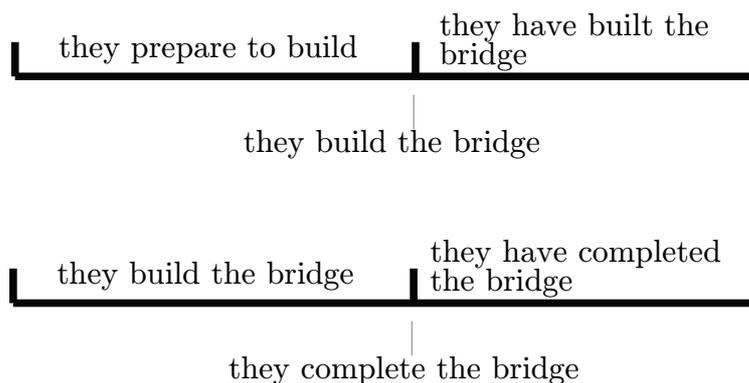


Figure 25 : Représentations temporelles chez Moens et Steedman

Nous ne garderons pas le principe de nucleus dans notre modèle, mais retiendrons que la durativité d'une situation n'est pas une propriété fondamentale de son image perceptive mais le résultat d'un calcul (voir les chapitres [7.2.6 Durativité](#) et [7.3.3 Opérations](#)).

Le modèle qui à notre préférence, même si nous n'en proposons pas d'implémentations, est celui de Gärdenfors (2014). Celui est compatible avec notre représentation du module perceptif et les principes énoncés par Ghadakpour (voir chapitre [6.3.1 Le système conceptuel](#), page 94) sur son fonctionnement.

Gärdenfors décrit les différentes valeurs aspectuelles comme des suites de vecteurs (ou paths) impliqués dans une situation. L'une des distinctions possibles est de savoir quel le vecteur est étendu ou ponctuel : Un path étendu peut être décomposé en une composition séquentielle de sous paths, tandis qu'un ponctuel ne peut pas l'être. En d'autres termes, on ne peut considérer qu'un moment. Cette distinction correspond en pratique à la distinction entre duratif et instantané, sauf que la dimension temporelle n'a pas besoin d'être mentionnée !

Une autre distinction du modèle de Gärdenfors est de savoir si le vecteur ou le path dispose d'un point de fin (endpoint) ou point fixe (fixpoint). **C'est uniquement cette distinction que nous utiliserons dans notre module perceptif sous le terme de télélicité.**

En prenant en compte ces deux distinctions, il est alors possible d'identifier les classes aspectuelles de Vendler.

Vector or path	<i>is extended</i>	<i>is punctual</i>
<i>has fixpoint</i>	Accomplishment (build a bridge)	Achievement (realize your error)
<i>has no fixpoint</i>	Activity (walk)	Semelfactive (knock)

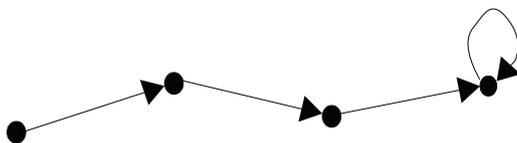
Tableau 5 : Classification of aspects, d'après Gärdenfors (2014)

Avec les représentations suivantes :

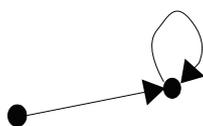
a) Un état (state) est un point fixe (fixpoint) : Absence de changement.



b) Un accomplissement (accomplishment) est une série de sous-événements (subevents) menant à un état final (télique).



c) Un achèvement (achievement) est comme un accomplissement (télique) mais consiste seulement en un instant.



d) Une activité (activity) est comme un accomplissement, sauf qu'elle ne se finit pas par un état final.



e) Une semelfactive (semelfactive) est une activité qui a lieu en une étape.



En repartant de la classification de Vendler, il reste l'attribut de dynamicité. Cette attribut semble nécessaire au calcul de l'aspect et de la temporalité, et ses valeurs sont particulièrement dépendantes de l'image perceptive. Nous pensons qu'il existe un moyen de la calculer (notamment afin d'éviter un accroissement de la complexité du modèle consécutif à l'addition d'un trait spécifique pour chaque image et expression ! Voir Ghadakpour pour la critique des traits) mais n'en proposerons pas de solution.

Au final, nous choisissons d'implémenter la **télicité** comme un switch (voir page 120) DANS le module perceptif et de l'attribuée à des images pré-construites. Ce choix simple va nous éviter d'avoir à choisir et implémenter un modèle de l'aktionsart, tout en montrant qu'il n'est pas nécessaire à notre module aspectuel d'avoir connaissance de comment la télicité est calculée.

La **dynamicité**, dont nous reconnaissons la pertinence pour l'interprétation finale, sera mise de coté puisqu'elle ne semble pas nécessaire pour le calcul de l'aspect "grammatical", celui dont notre module se charge.

La **durativité** (dans le sens de Vendler⁴⁰ !) nous l'avons déjà dit en citant Moens et Steedman, puis Gärdenfors, n'est pas nécessaire pour le calcul de l'aspect, et sans doute même complètement absente d'un module perceptif idéal.

6.3.3 L'ordonnement

La structure narrative implique un ordre chronologique, c'est-à-dire la capacité (mais pas

40 Nous reparlons de "durativité" plus tard dans le modèle mais dans un sens très différent.

l'obligation : la mémoire peut être incomplète !) d'ordonner des événements entre eux. Son importance pour le traitement du temps, et de la temporalité en particulier, est évidente.

6.3.3.1 Connaissances et causalité

Nous distinguons deux méthodes pour l'ordonnement des images perceptives :

La première est l'information mémorisée. Elle peut être utilisée dans le cadre de listes (les jours de la semaine par exemple) ou à propos d'événements historiques ou vécus (*Mozart est né avant Schubert, Marie était en Australie pendant l'hiver*, etc). On considère que la perception peut disposer de ces informations à propos des images perceptives : L'image d'un hiver particulier est simultanée avec l'image de son voyage en Australie. Nous proposons ici que le module perceptif ne dispose que de deux relations d'ordre partiel pour ordonner les images : la **précédence** et la **simultanéité**. La postériorité étant obtenue par symétrie avec la précédence. Qu'en est-il des relations plus complexes que l'on peut obtenir avec des modèles qui utilisent des intervalles de temps (Allen, 1984) ? Par exemple les chevauchements d'intervalles comme le recouvrement (125) ou la succession figure 27 ?

(125)



Figure 26 : Recouvrement d'intervalles

(126)



Figure 27 : Succession d'intervalles

Il se trouve que nous n'en avons pas besoin. Comment néanmoins distinguer les débuts et fins des situations ?

Nous proposons de traiter ces cas en **créant de nouvelles images** : S'il est nécessaire de distinguer le début d'une situation, ou sa fin, ou une étape particulière de son déroulement, on

considère que la perception disposera d'un ordonnancement entre les images de *début de A* et de *début de B*.

La seconde méthode d'ordonnancement des images perceptives est un **calcul** et ne se trouve (a priori) pas dans le module perceptif. C'est un module, que l'on suppose distinct, gérant la **causalité** qui fournira l'information d'ordonnancement. Pour reprendre l'exemple (91) du pont de Moens et Steedman l'ordre des situations est déduit de la connaissance ou de l'inférence de ce qui est la cause et la conséquence.

La gestion de la causalité sera mise de côté et ne fera pas l'objet d'un formalisme dans ce document, aussi nous nous contenterons d'utiliser les connaissances sur l'ordonnancement entre images du module perceptif.

D'un point de vue plus formel, comment représenter cet "ordonnancement entre images" ?

6.3.3.2 Des images relationnelles

Au début de notre recherche, nous avons considéré d'un côté des **images** et de l'autre des **"opérateurs"**. La nature de ces opérateurs était distincte des images : C'étaient des relations binaires (ils demandent à lier deux images entre elles) avec une valeur sémantique. Cette valeur était "précède" ou "est simultanée", mais pouvait être facilement étendue et exprimer des relations non-temporelles comme "est une partie de", "est", "appartient à la catégorie", etc. Cette structure est bien connue dans le domaine des ontologies sous le terme de triplets.

Le choix de construire un module perceptif sur le modèle des ontologies c'est révélé problématique à plusieurs égards.

Tout d'abord parce qu'aujourd'hui aucune ontologie n'est parvenue à simuler correctement l'organisation des connaissances d'un être humain. Celles-ci (Cyc, Babelnet, DBpedia, YAGO, etc.) ont de graves problèmes d'échelle : Une opération ou une recherche peut prendre plusieurs secondes, leur mise à jour (ajout de nouvelles informations) est encore plus lente, et surtout les ontologies ne parviennent jamais à être complètes (Fodor et al., 1980) puisqu'il semble toujours possible de trouver un lien absent qui sera pourtant évident pour un être humain (voir Ghadakpour, 2003 pour une critique des ontologies).

L'une des faiblesses des ontologies est l'arbitraire de la sémantique des liens : Travaillant sur le temps nous avons choisi d'avoir des liens topologiques ("précède", "est simultanée", "est une partie de", etc.). Mais une ontologie se focalisant sur les domaines de la connaissance humaine pourrait plutôt choisir des liens du type "appartient à la catégorie", "est un/e", etc. Une ontologie désirant représenter les liens de parenté aura les liens "père", "mère", "frère", "sœur", etc. Selon l'approche choisie la liste semble sans fin, et finalement particulièrement liée à l'expérience de chacun.

Nous aurions pu ne pas nous en soucier, et accepter de simplifier notre implémentation du module sémantique comme nous l'avons fait pour les **images**, mais un autre problème est apparu alors que nous mettions en place le module syntaxique et le mécanisme de **fusion** :

La fusion syntaxique demande deux éléments en entrée. Par parallélisme le module aspectuel et le module perceptif devaient eux aussi avoir deux éléments en entrée, et un seul en sortie. Une fusion à trois éléments (deux images et un lien par exemple) compliquait la chose et aurait contraint à imaginer des opérations de contournement ou à abandonner l'idée de fusions parallèles. Nous ne voulions surtout pas augmenter la complexité du modèle ni perdre le lien avec le module syntaxique. Que faire ?

Une solution se présenta qui évitait tous ses écueils et en plus permettait de se passer de la connaissance du fonctionnement du module perceptif :

Traiter les opérateurs relationnels portés par des mots comme *avant* ou *pendant* comme des **images**.

Ainsi *en* et *pendant* ont pour image perceptive l'image de **simultanéité**.

Nous simplifions notre vision du module perceptif (il n'y a plus de distinction entre le sens des liens et le sens des objets comme dans les ontologies ; (voir Hurford, 2003 pour l'approche en logique) en traitant *après* comme un opérateur unaire, comme le font les syntacticiens (Nous nous inspirons de Guéron (1993) pour l'arbre syntaxique de l'exemple (74) qui suit).

(74) *Marie est partie après le match.*

Une approche syntactique peut le laisser entendre. En syntaxe, les arbres syntaxiques peuvent être binaires, ce qui signifie qu'un adverbe ne peut s'associer directement qu'avec un seul élément. C'est l'ensemble qui pourra à son tour être associé à un autre niveau hiérarchique. Voici

l'arbre syntaxique de l'exemple :

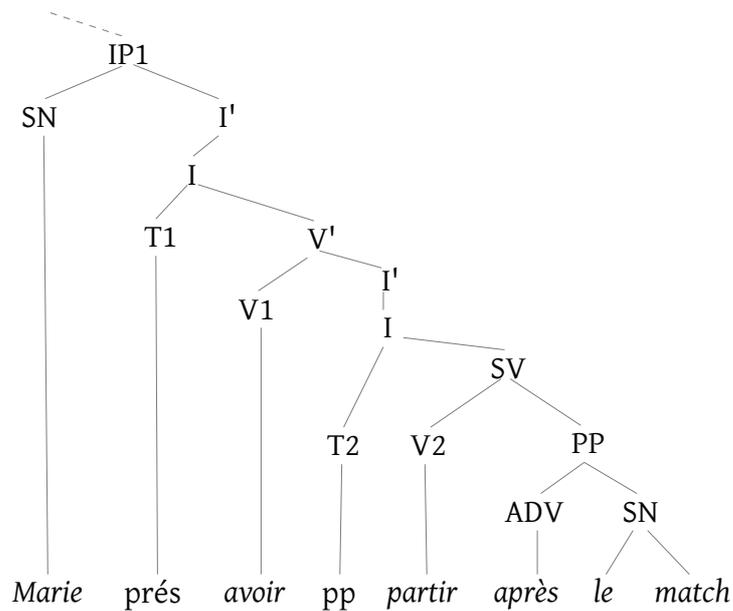


Figure 28 : Arbre syntaxique de (74)

après le match forme un syntagme prépositionnel, et ce n'est qu'au niveau hiérarchique supérieur, avec le syntagme verbal, que *après* trouve son second argument. Attention, cette approche suppose une "corrélacion directe" entre les liens syntaxiques et les liens sémantiques, ce qui est le cas dans notre modèle.

Nos fusions restaient binaires et on pouvait se passer d'utiliser des liens avec une valeur sémantique. Les opérateurs topologiques traditionnels (*avant*, *pendant*, etc.) devenaient des mots avec un comportement identique à tous les autres (on peut parler d'**opérateurs unaires**)! En plus de résoudre nos problèmes, cela nous permettait de traiter des énoncés du type *avant l'événement*, au lieu de s'attendre à des énoncés "complets" du type *Elle est arrivée avant l'événement*.

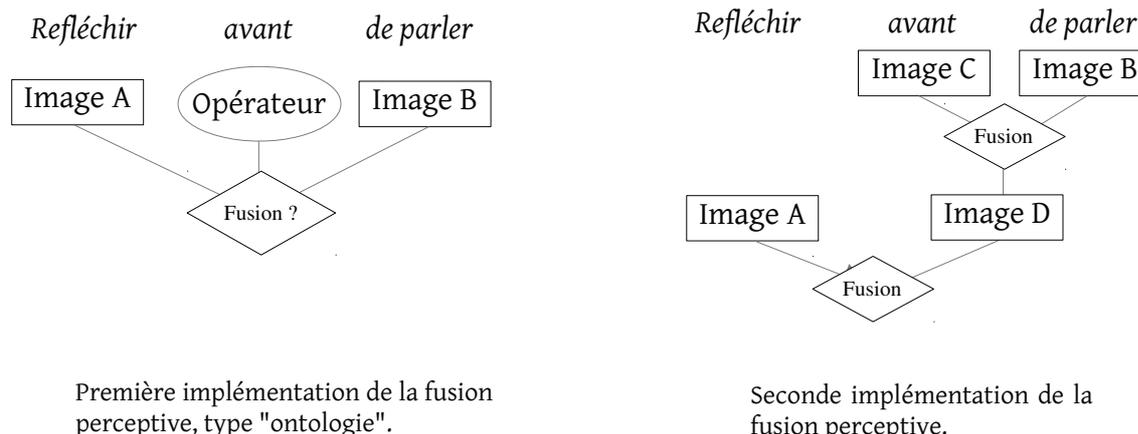


Figure 29 : Comparaison de deux types de fusions perceptives

Les opérateurs relationnels devenant des images, il est alors possible d'en créer de nouveaux sans devoir réviser tout le modèle. Les "opérateurs" sont acquis par l'expérience (voir Gerasymova, Spranger, & Beuls, 2012 pour l'acquisition de termes aspectuels pour le russe avec FCG). *Traverser les épreuves* ne demanderait pas de créer un nouvel opérateur de *traverser* pour être traité, ni *dépasser la date* un algorithme spécifique.

Le choix de placer l'information sémantique de ces "opérateurs relationnels" dans le temps permet d'être plus parcimonieux : Les informations fournies par le lexique sont peu nombreuses (elles se limitent pour nous aux cinq attributs d'une structure de taille fixe, que nous présentons dans le chapitre suivant), et celle fournies par la perception seront réutilisables dans des situations qui parlent d'espace (ou d'une autre "dimension" sémantique) plutôt que de temps.

6.3.3.3 Un calcul de l'ordonnement

Si ces "opérateurs relationnels" traditionnellement binaires sont réduits à des opérateurs unaires, comment le module perceptif se débrouille-t-il pour disposer des liens entre deux images (deux situations) ?

Nous considérerons ici que ces informations sont **calculées** lors de la fusion, parfois en faisant appel à des informations de causalité. Nous envisageons deux cas de figure :

- Le module perceptif reçoit lors de la fusion deux identifiants d'images (par exemple "*remise des médailles*" + "*après le marathon de paris*") et renvoie alors l'identifiant d'une nouvelle image valide, (pour nos exemples cela pourrait être une instanciation particulière de *remise des médailles* qui correspond à du vécu et pas à une image prototypique).
- Le module perceptif reçoit lors de la fusion un seul **identifiant** d'image (par exemple "*après le*

marathon de paris") mais renvoie néanmoins l'identifiant d'une image autorisée par le "contexte" (c'est-à-dire les images précédemment utilisées). Cette image sera soit identique à celle reçue, soit elle sera nouvelle (dans le cas de notre exemple, l'image renvoyée pourrait être *être épuisé, remise des médailles, lundi 7 Avril, nouvelle course*, etc.).

L'implémentation dans le module perceptif et le calcul opéré nous est inconnu, mais nous postulons des liens bidirectionnels entre les images et la possibilité de faire appel à des informations de causalité et des capacités déductives pour retrouver l'ordonnancement entre les images. Les possibles erreurs faites par les humains (Auriat, 1996) laissent penser que peu ou aucun de ces ordonnancements sont stockés statiquement ou font appel à des liens unidirectionnels. Nos exemples précédents avec *dépasser, traverser*, et plus souvent *avant, après*, etc. suggèrent un traitement unifié par un module **topologique**.

6.3.4 Résumé des propriétés du module perceptif

Nous plaçons dans le module perceptif toutes les données et traitements sémantiques qui n'appartiennent pas à notre module temporel. L'organisation et le contenu exact du module perceptif nous est inconnu, mais nous pouvons néanmoins relever quelques interactions précises entre module aspectuel et module perceptif :

- Le module aspectuel transmet au module perceptif l'indication d'une fusion de deux images en lui transmettant deux identifiants perceptuels. Il reçoit en réponse l'identifiant de l'image résultant de la fusion perceptive, ou une réponse négative en cas d'échec.

Les informations d'ordonnancement (précédence, simultanéité) sont obtenues (ou acceptées/refusées) lors de cet échange.

- Le module aspectuel demande la compatibilité des ordres de grandeurs de deux images perceptives. Il fournit pour cela les deux identifiants de ces images et obtient une réponse ternaire (oui, non, ne sais pas). Nous supposons que le module perceptif stocke ou calcule l'ordre de grandeur temporelle de chaque image (si ce n'est pas une durée relative à d'autres images, du type une 'heure' correspond environ à soixante 'minutes').

- Si le module aspectuel stocke des informations sémantiques non-perceptives (que nous allons détailler dans le chapitre suivant), il ne connaît pas les informations de dynamicité et de télélicité

d'une image. Ces informations sont soit stockées dans le module perceptif, soit (comme nous l'espérons) calculées à partir de propriétés/liens des images elle-mêmes. Par conséquent, le module aspectuel demande au module perceptif les informations de télélicité et de dynamicité, et reçoit des réponses binaire (oui/non).

7 Description du modèle

7.1 Vue d'ensemble

Nous avons dans le chapitre précédent, 6.3 Le module perceptif, présenté les informations nécessaires au traitement de l'aspect dépendant des images perceptives. Nous allons dans ce chapitre essentiellement décrire le module aspectuel et ses interactions avec les modules lexico-syntaxique et perceptif⁴¹.

Le système dans son ensemble fonctionne seulement en interprétation (bien qu'il ait été conçu pour s'adapter à un fonctionnement en production d'énoncés). Il reçoit en **entrée** des énoncés, pas forcément complets ou grammaticalement corrects.

Le système produit en **sortie** ce que nous appelons des arbres des fusions, qui correspondent à un ensemble d' AIS (chapitre 7.2 AIS : La Structure d'Information Aspectuelle) et de aMerge (chapitre 7.3 aMerge : La fusion sémantique (aspectuelle), page 129).

41 On suppose que le module aspectuel a des interactions avec d'autres modules sémantiques symboliques, par exemple pour connaître l'ordonnement de deux situations liées par un lien de causalité.

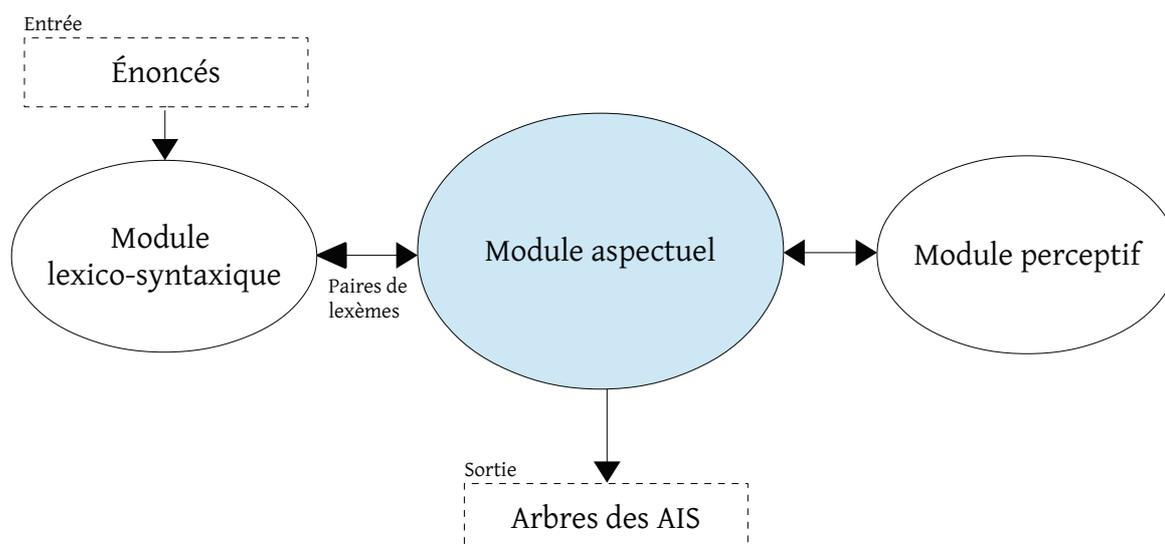


Figure 30 : Système global en mode interprétation

Les interactions entre le module lexico-syntaxique et le module aspectuel ont été décrites dans le chapitre [6.2 La fusion simultanée dans la syntaxe et la sémantique](#), page 90.

Les interactions entre le module perceptif et le module aspectuel ont été en partie introduites dans le chapitre [6.3.2 L'image perceptive](#), page 97.

Ces deux groupes d'interactions sont détaillés dans le chapitre [7.3.1 Interactions du module aspectuel](#), page 131.

Le cœur du modèle est le **module aspectuel** qui stocke et traite un seul type de structure sémantique symbolique : l'Aspectual Information Structure (AIS), décrite dans le chapitre [7.2 AIS : La Structure d'Information Aspectuelle](#), page 117.

Tous les traitements sur les AIS sont contenus dans le aMerge, ou fusion aspectuelle, décrit dans le chapitre [7.3 aMerge : La fusion sémantique \(aspectuelle\)](#), page 129.

Le module lexico-syntaxique a été implémenté en prolog et est un analyseur grammatical simple (les règles utilisées pour nos tests sont disponibles en annexe page 194).

Le module aspectuel va recevoir du module lexico-syntaxique des paires ordonnées de lemmes correspondant aux feuilles de l'arbre syntaxique de l'énoncé. Dans l'énoncé *Jean frappe la balle*, c'est la paire *la/balle* qui va d'abord être fournie, puis *frapper/la balle*, etc. (voir page 92).

Cette paire ordonnée de lemmes est l'unique élément demandé au module lexico-syntaxique : Tans que cette contrainte est respectée, le module aspectuel pourra communiquer avec n'importe

quelle modélisation et implémentation de la grammaire.

Les AIS respectifs pour chacun des lexèmes vont être identifiés et le module va déclencher la fusion (aMerge) de ces paires de AIS. Parallèlement à l'arbre syntaxique sera construit un "arbre des AIS" (voir la Figure 38, page 131, pour l'exemple *Jean frappe la balle*).

L'interprétation de cet arbre va nous permettre d'obtenir les informations aspectuelles finales de l'énoncé. Cette interprétation est décrite dans le chapitre [7.4 Lecture des arbres de AIS](#), page 148.

Nous proposons un exemple détaillé de traitement dans le chapitre [8.1 Exemples de traitements](#), page 151.

Les résultats de l'implémentation elle-même et quelques commentaires se retrouvent dans le chapitre [8.2 Implémentation](#), page 158.

7.2 AIS : La Structure d'Information Aspectuelle

7.2.1 Motivations

L'idée directrice de ce modèle a été de poursuivre le modèle topologique proposé par Ghadakpour (2003), modèle qui refuse une représentation d'un temps réaliste et une approche purement symbolique (comme la syntaxe) ou purement sub-symbolique (comme les réseaux de neurones, les modèles statistiques). Ghadakpour s'attaque à l'idée d'un modèle conceptuel symbolique, mais soutient l'idée d'une interaction entre "parties" symboliques et sub-symboliques.

Dans les propositions de Ghadakpour, seul le viewpoint était un attribut clairement identifié, et il n'existait ni structures ni formalisations d'un traitement du langage.

Le viewpoint est un type semblable - sinon identique - à la perfectivité (voir page 46). Il peut être soit une figure (perfectif, ou "moment"), soit un ground (imperfectif, ou "époque"). La figure, symbolisé par un rond "O" est la représentation d'une situation "vue dans son ensemble". Le ground, symbolisé par une ligne en pointillés " - - - - ", est la représentation d'une situation "vue depuis l'intérieur, sans connaissances des bornes".

Dans l'ébauche de modèle de Ghadakpour des paires de viewpoints s'enchaînent pour former des

"grilles temporelles". Chaque viewpoint était une représentation possible d'une situation, et chaque paire était reliée par un opérateur binaire ("simultané" ou "précède").

(127) *Il a peint la chambre pendant une heure*

L'énoncé (127) pouvait avoir une représentation graphique de cette forme :

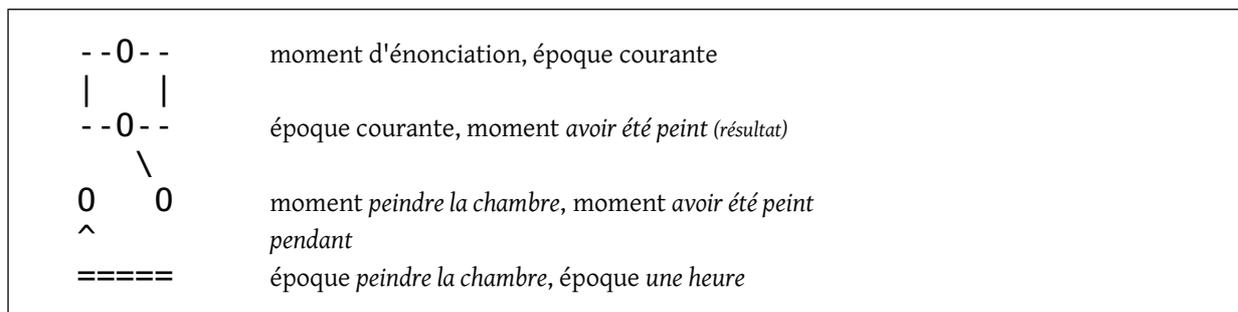


Figure 31 : Une "grille temporelle"

Les paires de viewpoints s'enchaînent en incluant à chaque fois un nouvel élément : Ce sont les "transitions". La relation et le type des viewpoints fournissent des informations aspectuelles et temporelles :

"--0--" par exemple est une simultanéité entre une figure et un ground, mais signifie en fait une inclusion : la figure est incluse dans le ground.

"0 0" comme autre exemple est la précédence d'une figure par rapport à une autre figure. Les deux viewpoints sont perçus comme achevés et ordonnés.

"====" enfin est une simultanéité entre deux grounds. Les deux viewpoints sont perçus comme inachevés.

Un viewpoint pour chaque paire est réutilisé dans la paire suivante, et peut par la même occasion changer de valeur (une figure en un ground par exemple, représente par le symbole "^", une figure qui reste une figure, représente par un trait "\").

Le système était simple et ne demandait quasiment rien comme structure de données (une variable de type binaire, deux relations topologiques, trois relations de transitions) alors qu'il pouvait exprimer des informations aspectuelles et temporelles beaucoup plus riches que le nombre de ces composants.

Mais ce modèle avait plusieurs zones d'ombre : les viewpoints représentaient des "situations", mais quelles sont-elles ? Les transitions portaient sur l'un des viewpoints de la paire, mais lequel choisir ? Comment représenter les répétitions ? Enfin, cette représentation ne parvenait pas à rendre compte de "détails", comme faire la différenciation entre *en 10 minutes* et *en 2010*.

Il manquait de nombreuses choses pour obtenir un modèle implémentable. "Nombreuses", mais le moins possible. Les axes directeurs pour le modèle de Ghadakpour qu'étaient la parcimonie et la plausibilité cognitive étaient toujours d'actualité et allaient encadrer nos travaux :

- Si deux viewpoints pouvaient fournir autant d'informations temporelles, il ne devait pas y avoir plus de deux ou trois autres éléments supplémentaires pour que leur combinatoire puisse exprimer la totalité des interprétations aspectuelles des langues naturelles.

- Ces éléments devaient être des variables, et en nombre fixe. On ne voulait pas se permettre de placer des algorithmes dans le lexique, ni de disposer de structures à taille variable : Si le modèle n'était pas contraint, rien n'empêcherait alors d'avoir des structures ou algorithmes de plus en plus vastes et puissants. Nous considérons que la richesse du langage provient de la combinatoire des éléments qui le composent, pas de quelques algorithmes attachés à des mots du lexique.

- Le traitement du temps fait partie du traitement du langage. Mais le traitement du langage, sa compréhension et sa production, fait appel à des informations qui vont au-delà de la syntaxe et du cadre commode du texte formaté. Ce que les sémanticiens appellent la pragmatique, c'est-à-dire la connaissance du monde, la prise en compte du contexte, la causalité, etc.

Notre modèle devait prendre en compte ces éléments, même si leur description et implémentation sont bien au-delà de nos possibilités, un ou plusieurs liens vers ces "modules pragmatiques" étaient des impératifs d'un modèle qui se veut cognitif.

De nos travaux (une version intermédiaire du modèle est brièvement décrite dans Munch & Dessalles, 2011) ont émergé une structure fixe (sa taille est toujours la même, il n'y a pas d'algorithmes qui y soient inclus) et non récursive (elle ne peut pas se contenir elle-même) : l'**Aspectual Information Structure (AIS)**, dont nous allons décrire les composants plus bas ; nous avons réussi à dégager un type de variable unique : le switch, et surtout un noyau algorithmique indépendant du lexique : le aMerge (voir page 129)

AIS

ImageID
Viewpoint
Determination
Duration
Multiplicity

7.2.2 Une variable binaire dynamique : Le switch

Avant de détailler les attributs de l'AIS, nous allons formaliser un nouveau type de variable utilisé dans l'AIS : Le switch (ou commutateur en français, mais nous préférons la forme anglaise pour la suite).

Le switch est une variable **binaire**. Un troisième état possible est celui d'être **non instancié**, que l'on pourrait interpréter comme "inconnu" ou "on ne sait pas encore". Cet état non instancié se comporte comme une variable prolog, c'est-à-dire qu'elle deviendra instanciée si on cherche à l'unifier avec une variable instanciée (nous détaillons "l'unification" page 133).

Les deux autres valeurs sont **instanciées** et peuvent être de sens différents selon la variable. On peut choisir les valeurs 1 et 0, ou "rouge" et "vert", ou oui et non, par exemple. Les attributs décrits par la suite ont chacun leur propre valeur possible.

Le switch prend la valeur de l'unificateur lors d'une unification avec un autre switch (voir page 133 le chapitre [7.3.2 Unification](#) pour l'explication de cette propriété).

Le switch a la particularité d'être **dynamique**. Cela signifie que sa valeur peut être changée au cours de la fusion, avec l'exécution d'autres opérations que l'unification (voir page 129).

7.2.3 ImageID

'imageID' est une référence (ou identifiant ; dans notre implémentation l'attribut se comporte comme une adresse ou un pointeur vers le module perceptif) d'une image perceptive. C'est elle qui permet de lier dans notre modèle un élément du lexique à son image perceptive. Ce n'est qu'une manière de trouver ce lien et il en existe assurément d'autres (dynamiquement suite à un calcul du module perceptif par exemple).

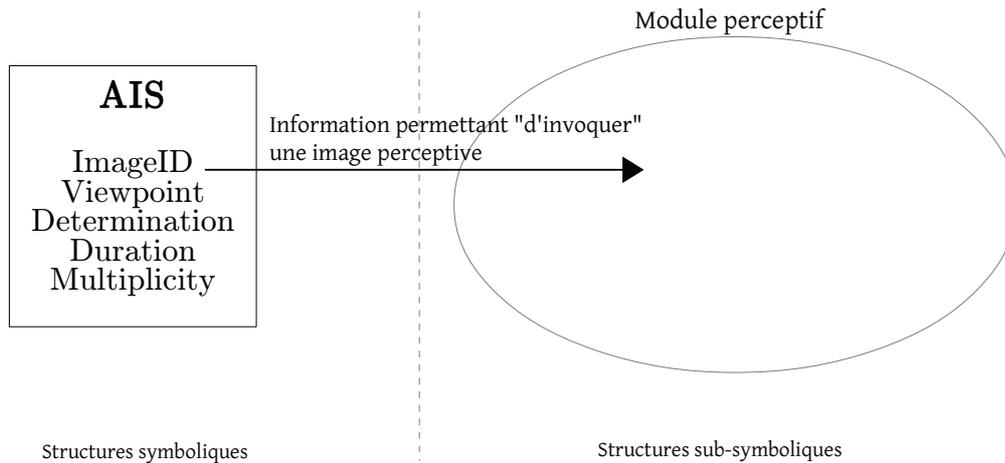


Figure 32 : Interface entre ImageID et image perceptive.

Au niveau de notre implémentation, une "image perceptive" est un atom prolog, c'est-à-dire un symbole.

Nous avons introduit sous formes de triplets des informations liées à ces atoms. Par exemple pour un atom *ronfler*, nous pourrions avoir les triplets suivants :

ronfler - has_magnitude_of - 2

ronfler - homogeneity - true

Nous le répétons, cette implémentation est ultra-simpliste et ne cherche pas à rendre compte du fonctionnement exact du module perceptif, mais permet uniquement au module aspectuel de recevoir des réponses possibles aux interrogations qu'il enverrait à un module perceptif "correctement" implémenté.

7.2.4 Viewpoint

Le 'viewpoint' est comparable à la notion de viewpoint chez Smith ou de perfectivité chez Filip (voir page 46 pour plus de détails) et fait directement référence au viewpoint de Ghadakpour.

Le 'viewpoint' est un switch. Il indique la manière de percevoir une situation : dans son ensemble ou en partie seulement.



Figure 33 : Point de vue et perspective

Les deux valeurs déterminées du 'viewpoint' sont le **ground**, qui correspond à la perception d'une partie seulement d'une situation (et dont on ne connaît pas les bornes), et la **figure**, qui correspond à la perception de l'ensemble d'une situation (dont on connaît les bornes).

Nous gardons les termes de Ghadakpour (plus populaires dans la psychologie de la perception que en linguistique) mais il serait possible de parler d'imperfectif pour ground et de perfectif pour figure.

Il existe évidemment un parallèle fort avec le point de vue spatial et nous pensons que notre module aspectuel et un module du traitement spatial font tous deux appel à un module dit "topologique".

Il semblerait que le point de vue peut être modifié par le "contenu" de la situation observée. En effet nous avons observé qu'une énumération d'objets est perçue comme une figure, quel que soit le 'viewpoint' de chacun des éléments. Par exemple (128) n'est pas correct (*manger du fromage* est un ground incompatible avec *en 5 minutes*) tandis que (129) est correct (*manger du comté et de la Fourme d'Ambert* est une figure) :

(128) # *il mangera du fromage en 5 minutes.*

(129) *il mangera du comté et de la Fourme d'Ambert en 5 minutes.*

La figure suivante propose des liens entre télélicité, viewpoint et perfectivité. Télélicité et perfectivité prenant les définitions discutées dans le chapitre [4.1 L'aspect](#), page 42 ; Le viewpoint prenant notre propre définition (de switch) :

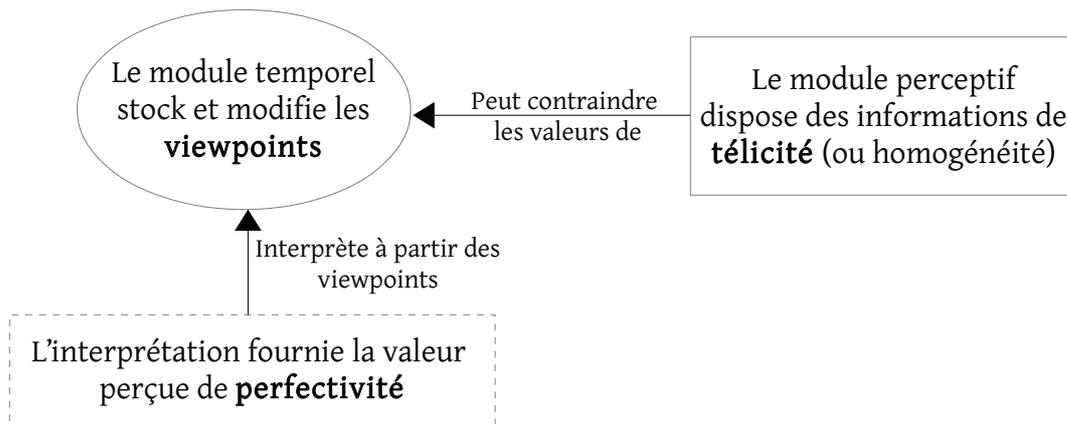


Figure 34 : Liens entre télécité, viewpoints et perfectivité.

Le 'viewpoint' peut être donné avec l'identification de l'élément lexical – l'imparfait ou la préposition *pendant* introduiront un ground ; le passé simple ou la préposition *en* introduiront une figure – ou être obtenu dynamiquement lors des sous-opérations de la fusion que nous détaillons dans le chapitre 7.3.3 Opérations.

7.2.5 Détermination temporelle

Nous avons déjà mentionné l'importance de ce que nous appelons ici "la détermination temporelle" en page 61. Pour rappel la détermination temporelle est ce qui distingue *ces dix minutes* de l'exemple (123) de *dix minutes* de l'exemple (122).

(122) *Elle a mangé du chocolat pendant dix minutes.*

(123) *Elle a mangé du chocolat pendant ces dix minutes.*

Ce phénomène est aussi connu sous le terme d'ancrage (anchoring) dans la littérature sur le temps (Mani & Schiffman, 2005; Pustejovsky et al., 2005). Nous avons choisi de garder le terme de 'determination' – en précisant qu'il est de nature temporelle – pour éviter une conséquence du sens d'ancrage que nous ne souhaitons pas : En parlant d'ancrage il serait nécessaire de parler de l'ancre à laquelle se rattache la situation, chose que ne voulons pas faire puisque n'avons pas de représentation discrète du temps, et que la durée déterminée n'est pas forcément ancrée de manière particulière à une autre situation.

Pour nous, être déterminé temporellement signifie que la ou les images perceptives dans le

contexte sont **distinguées** sans ambiguïtés des images prototypiques d'une unité syntaxique identique.

Par exemple, *dix minutes*, hors contexte et pris indépendamment sera undetermined parce qu'il n'est pas possible de distinguer leur image d'autres *dix minutes*. Par contre, si on fournit du contexte qui permettrait à la perception d'isoler et de distinguer l'une de ces images de l'autre, alors elles deviennent determined. Si je parle par exemple des *dix minutes avant la fin du match*, celles-ci seront distinguables de n'importe quelle autres *dix minutes* grâce au contexte.

(130) *Il fait froid en hiver.*

Dans (130), l'AIS de *hiver* est undetermined puisque nous ne pouvons (ou ne voulons) pas distinguer de quel hiver nous parlons. Cela peut aussi bien être celui de cette année que d'il y a mille ans.

(131) *Elle a appris à skier cet hiver.*

(132) *Elle a appris à skier en hiver.*

Dans (131), il n'y a pas de doute possible. Observons que (132) est ambigu : On ne sait pas si son apprentissage dure sur plusieurs hivers indéterminés (elle apprend peut-être toujours), ou si l'on parle d'un hiver en particulier. Les deux cas sont possibles, et c'est dans ce genre de situation que le module perceptif contraindra ou non la 'Determination'.

La 'determination' est un switch tout comme le 'viewpoint'. Il peut être renseigné avec le lexique ou obtenu lors de la fusion. Les deux valeurs instanciées possibles sont **determined** et **undetermined**.

À noter que la 'determination' nous permet d'éviter de catégoriser les "durées" des situations. Une durée comme *dix minutes* est représentée par une AIS avec un 'imageID' instancié et une 'determination' undetermined. On considère qu'une durée non déterminée n'est pas distinguable d'une autre durée (ou en tout cas nous n'avons aucun besoin de le faire). Voici quelques exemples d'énoncés ayant des AIS avec un 'imageID' instancié et determined : *dix minutes, trois nuits, en mangeant*.

Les situations au contraire sont distinguables les unes des autres et ont une 'determination' instanciée à determined. Voici quelques exemples : *l'année 2010, ces dix minutes là, hier, elle est arrivée en avion*.

7.2.6 Durativité

L'attribut de 'durativity' indique si l'image référencée par l' AIS a une valeur temporelle ou pas. Nous avons déjà abordé ce terme page 43 et allons en préciser ici le comportement.

La 'durativity' est un switch. Les deux valeurs déterminées possibles sont **durative** et **non-durative**. Tout comme les 'viewpoint' et 'determination', la 'durativity' peut être renseignée par le lexique et être contrainte par le module perceptif. Sa valeur peut aussi changer au cours du aMerge.

Le sens de durative ne signifie pas seulement, "qui a une durée" ou "qui a une existence temporelle", mais aussi et surtout "dont l'existence temporelle est prise en compte", ou en faisant un parallèle avec le 'viewpoint' : "où l'on perçoit l'existence temporelle".

Non durative a donc une signification en opposition à durative : "Dont on ne perçoit pas l'existence temporelle". Cela peut inclure tout aussi bien des images atemporelles comme des couleurs, des concepts ou des relations (*avant, pendant*), mais aussi des images qui peuvent avoir une existence temporelle mais qui sont perçues de manière "logique".

Voici deux exemples qui illustrent cette distinction :

(133) *Il a couru un marathon, lui !*

(134) *Il a couru un marathon en trois heures et demie.*

Dans (133), *courir un marathon* n'est pas **perçu** comme une situation temporelle. L'information fournie ici est qu'il a fini le marathon, ou au moins qu'il a essayé. *Courir un marathon* est perçu d'un point de vue logique. Dans (134) au contraire l'information temporelle (la 'durativity') de *courir un marathon* est bien présente. La distinction de ce que l'on veut exprimer entre ici en jeu et fait appelle à la notion de prédication (voir chapitre [7.3.3.2 Focus-out et "prédication" temporelle](#)).

On rejoint ici la position de Moens & Steedman (voir page 101) en ayant suivi une approche différente.

7.2.7 Multiplicity

La 'multiplicity' est l'information de pluralité des situations (Talmy parle de "plexity" (Talmy, 2000)). La 'multiplicity' est dans notre implémentation un switch. Les deux valeurs déterminées possibles sont **multiple** et **single**. Contrairement à la 'durativity', la 'multiplicity' peut-être apportée directement par les éléments du lexique, par exemple dans *plusieurs fois* ou le déterminant *des*.

La 'multiplicity' peut aussi être obtenue après un calcul comme c'est le cas pour l'exemple (58) (voir chapitre [7.3.3.3 Repetition](#)).

(58) *L'an dernier, Marie mangeait à la cantine.*

On remarquera que l'information de pluriel n'est absolument pas limitée au temps et aux situations ! Nous avons pourtant choisi dans notre implémentation de distinguer le pluriel "classique" du pluriel appliqué au temps (pluralité des situations), en lui donnant le nom de *multiplicity* plutôt que *plurality*. Nous sommes conscients qu'il serait bien plus élégant et parcimonieux de n'avoir qu'une seule variable et de ne pas distinguer pluriel temporel et pluriel non-temporel. Ce choix a surtout été fait pour faciliter les recherches et éviter de devoir aborder en détails les études sur la détermination, la quantification, l'énumération, etc.

7.2.8 Les AIS dans le modèle

Nous considérons que TOUS les éléments du lexique (ou lemmes) disposent de leur propre AIS. Les attributs de l'AIS ne sont pas forcément instanciés et il est tout à fait possible d'avoir des AIS partiellement renseignés. Nous noterons par un point d'interrogation "?" les variables indéterminées (identifiant ou pointeurs). Les lemmes qui n'ont ni image ni ne portent d'informations "temporelles" disposeront tout de même d'une AIS entièrement non-instanciée. Voici quelques exemples :

AIS Jour	
ImageID	i_day
Viewpoint	?
Determination	?
Duration	?
Multiplicity	?

AIS imparfait fr.	
ImageID	?
Viewpoint	ground
Determination	?
Duration	?
Multiplicity	?

AIS présent simple fr.	
ImageID	?
Viewpoint	?
Determination	?
Duration	?
Multiplicity	?

AIS simple past anglais	
ImageID	?
Viewpoint	figure
Determination	?
Duration	?
Multiplicity	?

AIS pendant	
ImageID	i_simult
Viewpoint	ground
Determination	?
Duration	?
Multiplicity	?

AIS en	
ImageID	i_simult
Viewpoint	figure
Determination	?
Duration	?
Multiplicity	?

AIS dix minutes	
ImageID	i_10min
Viewpoint	?
Determination	?
Duration	durative
Multiplicity	?

AIS 2010	
ImageID	i_2010
Viewpoint	?
Determination	determined
Duration	durative
Multiplicity	?

Figure 35 Exemples d'AIS

Les 'imageIDs' sont donnés à titre indicatif et dépendent surtout de l'implémentation du module perceptif. Les temps (comme l'imparfait) incluent les conjugaisons. De plus, les variables de temporalités ont été volontairement écartées de cette présentation parce que nous ne nous en servons pas et surtout parce que notre implémentation de la temporalité a été insuffisamment testée (voir chapitre 10.2 la temporalité et l'aspect pour une discussion à ce sujet).

Dans la suite du document, nous pourrons représenter les AIS dans le texte en mettant des crochets (" []") à la fin de l'énoncé et en lui assignant une nouvelle police. Par exemple, l'AIS de *en* dans le texte est celui-ci : **en []**. Sa représentation graphique avec ses valeurs aura cette forme :

	imageID	viewpoint	determination	durativity	multiplicity
<i>En</i>	i_simult	figure	?	?	?

Ci-dessous la création de l'AIS de l'imparfait dans l'implémentation prolog. 'vp' correspond au viewpoint, tandis que le dernier underscore correspond aux valeurs non-instanciées. 'dsem_unit' décrit une unité sémantique dynamique, le préfixe du premier argument, 's_fr_' décrit un identifiant sémantique pour la langue française.

```
assert(dsem_unit(s_fr_tense_imparfait, [vp:g|_]))
```

Insistons sur le fait que les informations de l'AIS et celles fournies par le module perceptif sont les seules nécessaires au traitement de l'aspect. Nous allons voir que l'AIS est uniquement composée d'une référence ('imageID', assimilable à une "adresse") et de quatre switches. C'est une structure de taille fixe, non récursive et sans algorithmes.

7.2.9 Détails d'implémentation

Pour rappel 'l'imageID' est un atom. Son nom sera de la forme "i_" + nom de l'image perceptive (qui est elle-même traitée dans l'implémentation comme un atom).

Le mot pomme aurait par exemple une 'imageID' de la forme "i_pomme", sachant qu'il existerait dans notre implémentation du module perceptif une image "pomme" qui correspondrait à l'image perceptive prototypique d'une pomme chez un individu.

Les quatre autres attributs, 'viewpoint', 'determination', 'durativity', 'multiplicity', sont des switches. Dans notre implémentation prolog, les deux valeurs possibles pour chacun sont des atoms, tandis que la valeur "non instanciée" correspond exactement au comportement natif d'une variable non instanciée de prolog.

7.3 aMerge : La fusion sémantique (aspectuelle)

Nous avons postulé le parallélisme des fusions syntaxique, sémantique et perceptive.

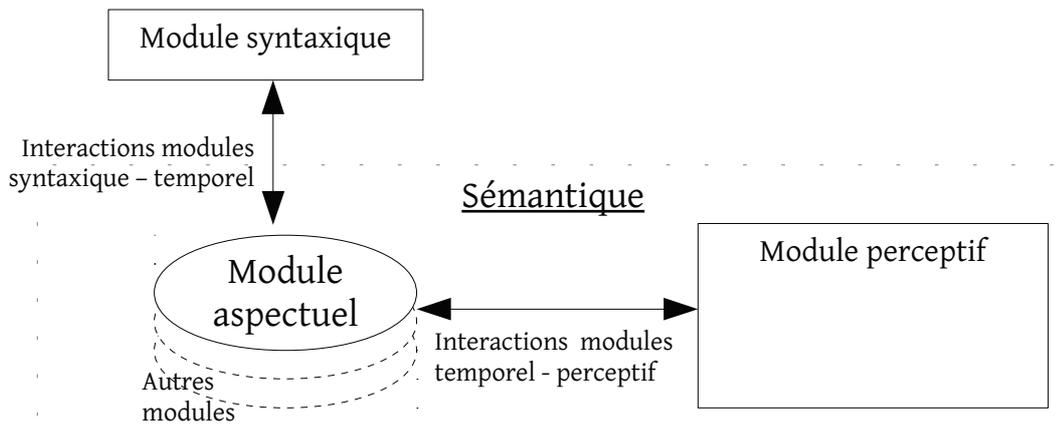


Figure 36 : Interactions entre les modules syntaxique, aspectuel et perceptif

La fusion syntaxique a été rapidement introduite dans le chapitre [6.2 La fusion simultanée dans la syntaxe et la sémantique](#) tandis que notre implémentation (une large simplification de ce qu'elle devrait être) de la fusion perceptive a été présentée dans le chapitre [6.3.2.1 La fusion perceptive](#).

Nous allons maintenant aborder le cœur de notre modèle : la fusion aspectuelle, que nous appellerons le **aMerge**, qui consiste essentiellement à fusionner deux AIS. La fusion aspectuelle consiste grossièrement en un mécanisme d'**unification** (dans le sens logique) et en une série d'**opérations**. Selon l'ordre de l'implémentation ces opérations peuvent être perçues comme des opérations de "résolution" de conflits (c'est le terme utilisé par Gosselin), si l'unification est en premier ; ou des opérations "systématiques", si celles-ci sont réalisées en premier.

La fusion aspectuelle est la fusion de deux AIS, l'une dominante, l'autre complémentaire (l'ordre étant fourni par la syntaxe). Son résultat est une unique AIS.

Ce mécanisme va permettre d'avoir un système non-monotone puisqu'il est possible de détruire les AIS fusionnées et d'éviter un accroissement non contrôlé du nombre de structures.

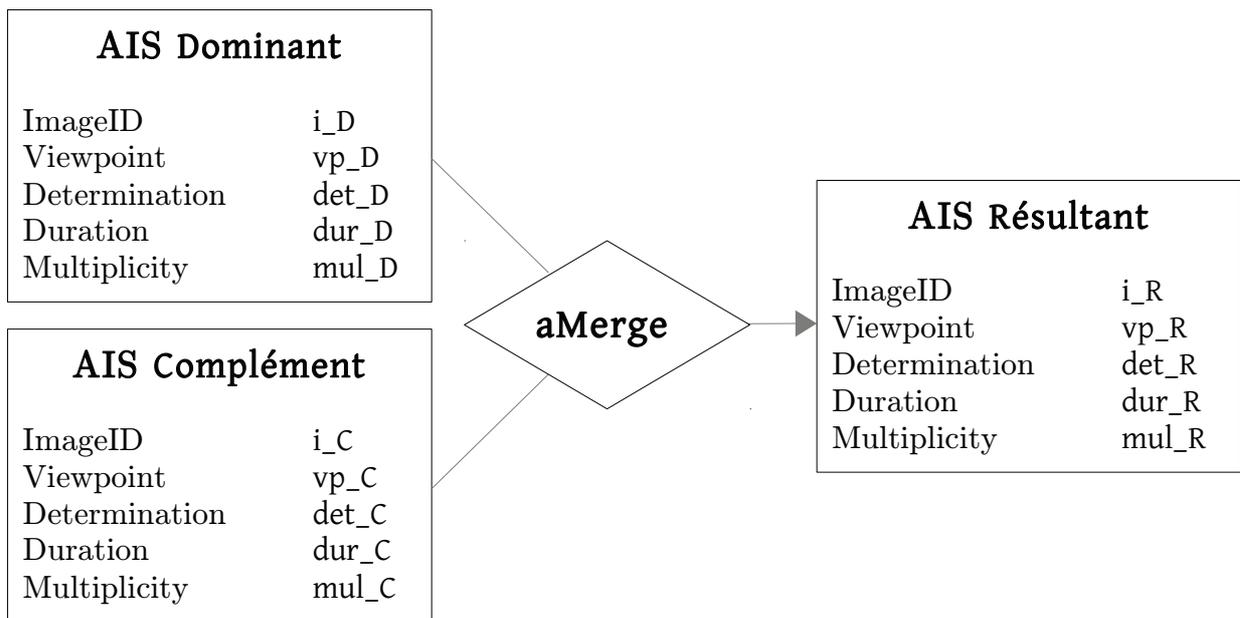


Figure 37 : Entrée et sortie du aMerge

Rappelons que le aMerge est guidé par la fusion syntaxique (page 90). Ci-dessous l'ordre des fusions des AIS de l'énoncé (117) selon la théorie syntaxique du programme minimaliste. **Les rectangles sont des AIS, les losanges sont des aMerge** (que nous détaillons dans le chapitre 7.3 aMerge : La fusion sémantique (aspectuelle), page 129). Les détails de chaque étape sont représentés dans les chapitres suivants.

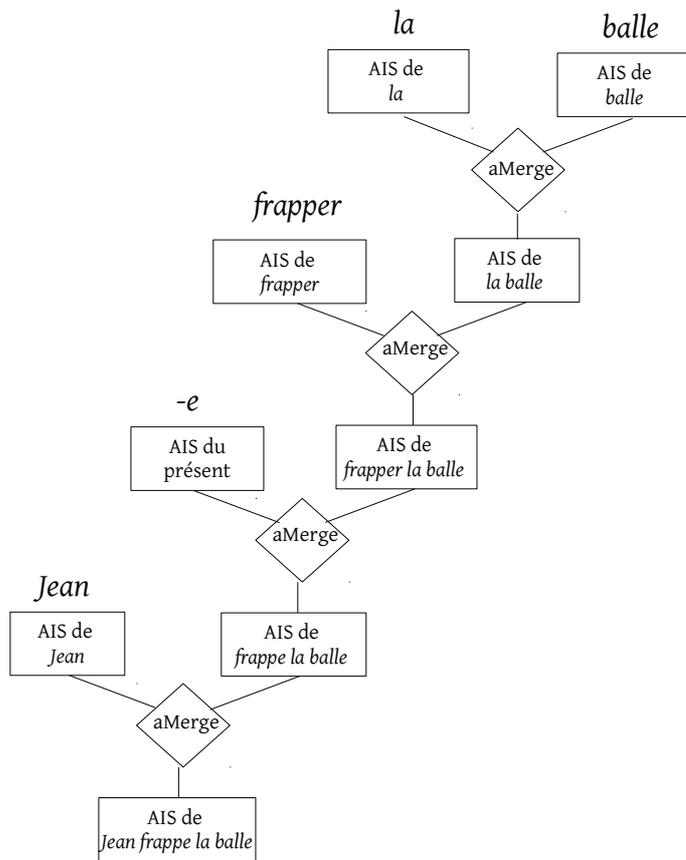


Figure 38 : Arbre des AIS de *Jean frappe la balle*.

7.3.1 Interactions du module aspectuel

Observons le comportement externe du module aspectuel : Il dispose de contacts avec le module syntaxique et le module perceptif (qui nous le rappelons est une simplification d'un ensemble de traitements sémantiques probablement indépendants, et donc modulaires, et des connaissances de la mémoire épisodique).

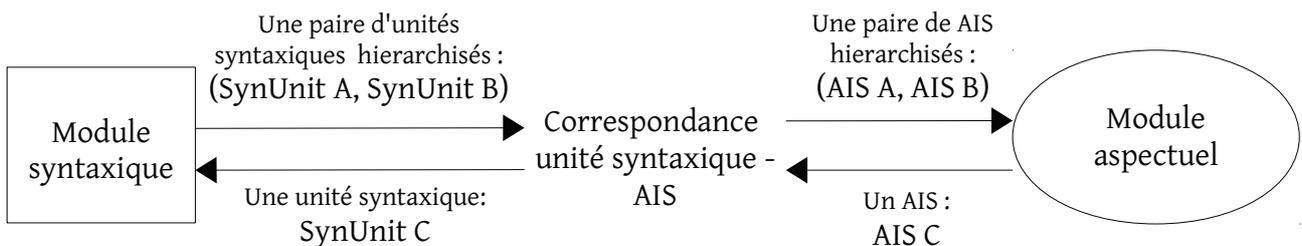


Figure 39 : Interface entre module syntaxique et temporel, en interprétation

Lors de la fusion, l'interaction entre les modules syntaxique et aspectuel demande de trouver une correspondance entre des unités syntaxiques et leurs AIS. Notre définition d'unité syntaxique est la suivante : Une unité syntaxique peut aussi bien être un lemme qu'un groupe hiérarchisé d'unités syntaxiques (un nom et son adjectif, *balle rouge* par exemple, sont individuellement des unités syntaxiques, mais leur groupement est lui aussi une unité syntaxique). En interprétation le module syntaxique envoie une paire hiérarchisée d'unités syntaxiques et reçoit une nouvelle unité syntaxique.

Le lemme *pendant* aura par exemple une correspondance avec le AIS suivant :

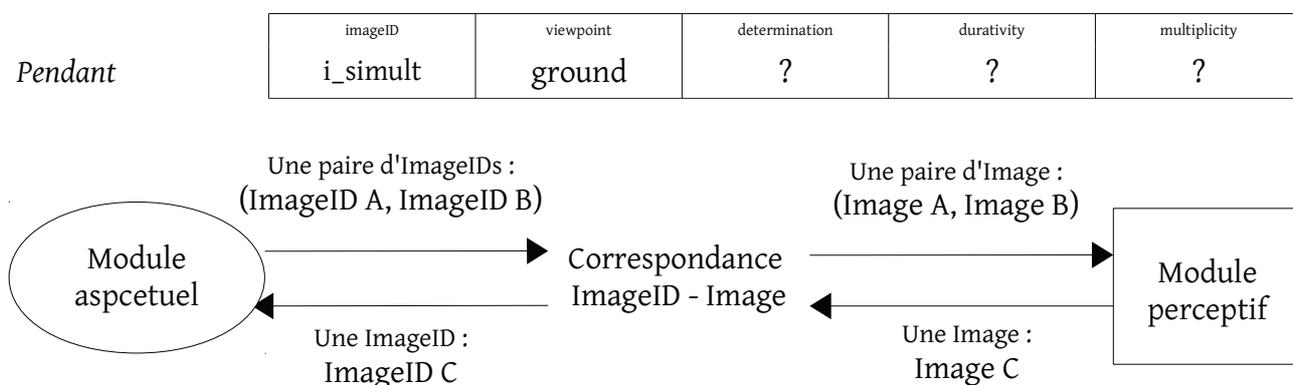


Figure 40 : Interface entre module syntaxique et temporel, en interprétation

Lors de la fusion, l'interaction entre les modules aspectuel et perceptif demande de trouver une correspondance entre des 'imageIDs' et des images (perceptives). En interprétation le module aspectuel envoie une paire 'd'imageID' et reçoit une nouvelle 'imageID' en réponse.

Dans notre implémentation, nous nous contentons de faire de la concaténation des atoms prolog correspondants aux 'imageIDs' : L'envoi des 'imageID' "i_balle" et "i_rouge" conduira à la réception de l'imageID "i_balle_rouge".

Les modules aspectuel et perceptif disposent d'autres interactions (question sur le 'viewpoint' possible, en lien avec la télécité de l'image perceptive), sous la forme de questions fermées et de réponses oui/non (par exemple : "cette image X peut-elle être une figure (être telique) ?"). Ces interactions sont particulières aux opérations de la fusion aspectuelle et nous en reparlerons à ce moment (chapitre [7.3.3 Opérations](#)).

7.3.2 Unification

En logique, l'unification de deux termes t_1 et t_2 consiste à trouver un troisième terme t tel que l'on puisse passer de t à t_1 et t_2 en instanciant certaines variables. t est alors appelé l'unificateur. Dans notre situation l'unification se fait directement sur des switches : les attributs des AIS, ce qui permet de réduire le nombre de cas possibles à quatre. Dans les exemples, le symbole "+" signifie "essaie d'être unifié avec", le symbole "→" signifie "entraîne", et on pose comme valeurs instanciables "A" et "B" :

1) Les deux switches sont non instanciés. L'unificateur peut alors prendre n'importe laquelle des valeurs possibles de l'attribut (les deux valeurs instanciées ou rester non instancié).

Si $t_1 = ?$ et $t_2 = ?$, alors $t_1 + t_2 \rightarrow t = ?$ OU $t = A$ OU $t = B$

2) Un switch est instancié et l'autre non. L'unificateur prend alors la valeur du switch instancié.

Si $t_1 = A$ et $t_2 = ?$, alors $t_1 + t_2 \rightarrow t = A$

3) Les deux switches sont instanciés avec une valeur identique. L'unificateur prend alors cette même valeur.

Si $t_1 = A$ et $t_2 = A$, alors $t_1 + t_2 \rightarrow t = A$

4) Les deux switches sont instanciés avec une valeur différentes. L'unification échoue.

Si $t_1 = A$ et $t_2 = B$, alors $t_1 + t_2 \rightarrow$ échec

Les switches non-instanciés prennent comme nouvelle(s) valeur(s) celle(s) de l'unificateur.

La valeur de l'unificateur est ensuite utilisée pour la valeur de l'attribut dans l'AIS résultante.

Nous avons un type de variable binaire simple : le switch, une variable d'adressage ('imageID'), et le principe d'unification. Les unifications de chaque attribut ont des comportements contraints liés au module perceptif.

7.3.2.1 Unifications avec contraintes perceptives

Les attributs de l'AIS 'imageID', 'viewpoint' et 'determination' partagent une même propriété : Au moment de l'unification, le module aspectuel demande au module perceptif de **valider** les valeurs

en fonction des 'imageIDs'. Cette validation fonctionne de manière différente pour les trois attributs.

'ImageID'

L'unification des 'imageIDs' se fait de manière classique (comme expliqué dans le chapitre [7.3.2 Unification](#)) dans les cas 1, 2. Les cas 3 (rares) et 4 sont différents puisqu'ils ne sont possibles qu'après l'interrogation du module perceptif. Ils sont assez particuliers puisque le résultat **de l'unification des 'imageID'** est en fait l'identifiant du résultat **de la fusion perceptive**. Nous donnons ici des exemples (E et F sont les identifiants de nouvelles images du module perceptif) :

Si $t_1 = A$ et $t_2 = A$, alors $t_1 + t_2 \rightarrow t = E$

Si $t_1 = A$ et $t_2 = B$, alors $t_1 + t_2 \rightarrow t = F$

Au niveau du module perceptif, cela revient par exemple à fusionner l'Image de *manger* (dans son sens prototypique) avec l'Image de *cantine* (prototypique ou pas selon le contexte) et d'obtenir l'image de *manger à la cantine* qui correspondrait par exemple à *manger à une table avec des collègues ...*

Le module perceptif renvoie alors un identifiant de cette nouvelle image (nous présentons cette fusion dans le chapitre [6.3.2 L'image perceptive](#)).

Dans notre implémentation, on se contentera de concaténer les atoms des images fusionnées. Ainsi l'image de *manger* et l'image de *cantine* donneront "repas_cantine". Les informations sémantiques qui sont attachées aux images (voir chapitre suivant) ne sont pas fusionnées : nous simplifions en les ajoutant directement dans notre implémentation du module perceptif pour chacun des résultats des images fusionnées rencontrées⁴².

Voici un extrait du code prolog pour la création d'une nouvelle image :

```
%% Nouvel atom
new_atom(Atom1, Atom2, Atom_result) :-
    atom_concat(Atom1, '_', Atom1b),
    atom_concat(Atom1b, Atom2, Atom2b),
    %% génération du nouveau symbole
    gensym(Atom2b, Atom_result).
```

'Viewpoint' et les contraintes de télécité

⁴² Cette sorte de "tricherie" ne vise qu'à simplifier notre système. Elle ne remet pas en cause la possibilité d'avoir une fusion perceptive automatique.

L'unification des 'viewpoints' (figure et ground) se fait de manière classique pour tout les cas. Il existe néanmoins un appel au module perceptif parallèle à l'unification classique, afin d'**accepter** ou non les valeurs des 'viewpoints'.

Les contraintes sont données par le module perceptif et prennent en compte la télélicité (voir page 48) des images perceptives qui ont été fusionnées. Au moment de l'unification le module aspectuel demande au module perceptif son accord. On peut formaliser la question posée de cette manière : "Est-ce que l'Image identifiée (par 'l'imageID') peut être ou devenir télélique ?".

- Si oui, alors le 'viewpoint' de l'AIS correspondante est compatible avec une figure.

- Si non, alors le 'viewpoint' de l'AIS correspondante est compatible avec un ground.

Par exemple :

(135) ?? *Il ronfle en deux heures.*

Dans l'énoncé (135), l'unification entre les AIS de *ronfle* et *en deux heures* va donner une figure. Le module perceptif à qui l'on pose la question sur la télélicité de l'image de *ronfler* répondra non, conduisant à un échec. Si la télélicité est donnée par le module perceptif, cela signifie que nous ne connaissons pas la manière dont la valeur est calculée ou stockée. Nous renvoyons aux Limites et Perspectives pour une discussion à ce sujet.

Notons qu'il existe une interprétation acceptable de (135) : Si la personne s'endort mais ne commence à ronfler que au bout de deux heures (*en* prend alors le sens de *dans*). Nous présentons ce mécanisme dans le chapitre 7.3.3.5 Shift.

On constate que la télélicité - fournie par le module perceptif - et le 'Viewpoint' - fourni par la syntaxe ou obtenu dynamiquement - sont très liés, mais pas identiques ! Nous évitons ainsi la critique faite au modèle de de Swart (page 82), qui est de confondre les deux et de perdre ces informations.

'Determination' et les contraintes d'instanciation

L'unification des 'determination' se fait de manière classique pour tout les cas. Il existe néanmoins un appel au module perceptif (tout comme pour le 'viewpoint'), qui précède l'unification classique, afin de **contraindre** les valeurs des 'determination'.

La valeur determined peut provenir du module perceptif plutôt que d'être le fruit de l'unification, si le module perceptif considère que l'image de l'AIS qui est en train d'être fusionnée peut être clairement distinguée. Nous nous étendons un peu plus sur la "distinction" dans le chapitre [7.2.5 Détermination temporelle](#).

Si une discussion est en cours à propos de la dernière semaine de vacances d'un ami, la 'determination' de *Lundi* est determined, puisque le module perceptif est certain que ce lundi est celui de la semaine de vacances, et pas d'un lundi générique.

7.3.2.2 Unification sans contraintes

La 'multiplicity' est particulièrement flexible lors de l'unification. Les valeurs des AIS, si elles contiennent toutes les deux des 'ImageIDs', deviennent non instanciées avant même l'unification. Cela signifie qu'elles se trouvent toujours dans les cas 1) et 2) listés au début de 7.3.2 Unification.

(136) *Elle a éternué plusieurs fois pendant le spectacle.*

Dans l'exemple (136), le modèle a réussi à unifier les AIS de *éternuer plusieurs fois* et *pendant le spectacle*, alors même que l'un est multiple et l'autre unique (*le spectacle*).

L'attribut de 'multiplicity' est pris en compte par le module perceptif, ce qui laisse penser que dans un système plus vaste (avec plus de modules sémantiques), les switches de 'multiplicity' et 'durativity' sont soumis d'une manière ou d'une autre à une contrainte du module perceptif, mais que dans notre cas ils peuvent se comporter avec moins de contraintes que les autres switches.

7.3.3 Opérations

Nous avons vu que les unifications pouvaient échouer (page 133). Dans ce cas, la fusion n'échoue pas forcément et il existe des mécanismes qui permettent de "corriger" ou "prévenir" l'échec. Les opérations peuvent avoir lieu avant l'unification (on parle d'opérations systématiques) ou après l'unification (on parle d'opérations de récupération). L'ordre n'est pas important et c'est essentiellement un choix d'implémentation. Nous choisirons de présenter les opérations comme systématiques pour la suite du document.

Les opérations ne s'appliquent que sur les syntagmes (les résultats de fusions). Cela signifie que l'on ne peut pas appliquer d'opération sur *en* mais plutôt sur *en 2010*. Ce comportement est compréhensible puisque le module aspectuel n'est a priori sollicité que lors d'une opération de fusion.

Nous avons identifié deux couples d'opérations qui peuvent s'appliquer l'un après l'autre : focus-in/focus-out forment le premier couple et repetition/slice forment le second. Cela signifie qu'il est possible d'appliquer, par exemple, d'abord une opération de focus-in, puis un slice. Une cinquième opération a été identifiée : le shift, ou déplacement.

Note : Lorsque nous utilisons le terme "instanciable", nous voulons dire que l'attribut est déjà instancié à la valeur indiquée OU qu'il est non instancié.

7.3.3.1 Focus-in

Le focus-in est une opération qui consiste à "se rapprocher" de la situation, ou à "rentrer dans" celle-ci. C'est une opération dont l'utilité est de passer d'un 'viewpoint' figure vers un ground.

L'opération n'est pas libre et il y a des contraintes à son application :

- 1) l'imageID est instanciée.
- 2) le 'viewpoint' est instanciable en une figure.
- 3) la 'determination' est instanciable en determined.

Si ces contraintes sont remplies, alors l'opération de focus-in peut s'appliquer et a les effets suivants sur l' AIS concernée :

- I) Le 'viewpoint' devient un ground.
- II) La 'determination' devient determined.

l' AIS de 2010 (l'année) est

	imageID	viewpoint	determination	durativity	multiplicity
2010	i_2010	?	determined	?	?

Ce qui signifie que le 'viewpoint' peut être une figure ou un ground. Si on prend le AIS de *en 2010*, sa AIS est

en 2010

imageID	viewpoint	determination	durativity	multiplicity
i_2010	figure	determined	?	?

Son 'viewpoint' est instancié et sans opérations il n'est pas possible de transformer l'AIS en un ground. Le focus-in rend la chose possible. Toutes les conditions étant remplies, le modèle peut "zoomer" sur *en_2010* [].

On remarquera que le focus-in n'est pas applicable sur

en dix minutes

imageID	viewpoint	determination	durativity	multiplicity
i_10min	figure	undetermined	?	?

à cause de la valeur undetermined.

7.3.3.2 Focus-out et "prédication" temporelle

Le focus-out est une opération qui fait l'inverse du focus-in : On transforme un ground (situation perçue "de l'intérieur") en une figure ("vue d'ensemble").

Attention, le focus-out ne se contente pas de transformer le 'viewpoint', mais agit sur la durativité, c'est-à-dire la manière de percevoir la situation. Nous avons remarqué pendant nos tests que si une opération de focus-out s'appliquait, alors l'image de l'AIS concernée **perdait** sa valeur temporelle, comme si l'information temporelle n'était plus pertinente, et que son unité syntaxique devenait le **prédicat** (au sens logique) de l'énoncé. Nous avons déjà parlé de ce phénomène dans le chapitre 4.4 Attitude propositionnelle.

Un indice dans l'interprétation de cette prédication temporelle est de comprendre qu'un événement exceptionnel, un record ou le viol d'une interdiction par exemple, avait eu lieu.

Reconsidérons l'énoncé (89) :

(89) *Marie a bu du champagne l'année dernière.*

Dans une interprétation on comprend que *Marie* a fait quelque chose qu'elle n'aurait pas dû ou pu faire : *boire du champagne* (On a envie de dire que cette année elle n'en a pas bu).

On pourrait nous rétorquer que cette interprétation est une déduction, liée à notre connaissance du monde. Mais le même phénomène apparaît si on remplace *boire du champagne* par *schtroumpher*, alors même que nous n'avons pas de connaissance du sens de ce verbe. La seule contrainte est de s'imaginer que *schtroumpher* prend moins d'un an.

Il semblerait que la "prédication temporelle" soit une conséquence de l'opération de focus-out. Attention cependant, le terme de "prédication" a plusieurs sens selon le domaine. Nous le comprenons ici dans son sens logique, mais même ainsi la "prédication temporelle" n'est pas uniquement une "prédication".

La prédication (pas forcément temporelle) dans le langage permet de mettre l'accent sur quasiment n'importe quel élément d'un énoncé. Pour reprendre l'exemple (89), cette prédication non temporelle peut porter sur *Marie* (c'était elle et pas *Jean*) ou sur *l'année dernière* (c'était *l'année dernière* et pas *il y a deux ans*) par exemple. On observe que dans ces cas (les interprétations ne sont pas évidentes...) la dimension temporelle ('durativity' dans notre modèle) de *boire du champagne* disparaît alors, mais qu'il faut alors faire appel au contexte pour que la phrase soit cohérente (l'énoncé doit être alors un écho ou une correction à une phrase similaire de la discussion).

Voici les contraintes d'application de l'opération de focus-out :

- 1) 'imageID' est instanciée.
- 2) Le 'viewpoint' est instanciable en un ground.
- 3) La 'durativity' est instanciable en durative.
- 4) L'opération de focus-out n'a pas déjà été utilisée dans la phrase.

Le point (4) est un ajout après observations : Une seule prédication est possible par phrase. Nous considérons qu'un module spécifique (un module "logique"?) est capable de stocker cette information, aussi nous nous contentons de la stocker hors d'une AIS dans une variable binaire.

Si toutes les contraintes sont remplies, l'opération de focus-out peut s'appliquer :

- I) Le 'viewpoint' devient une figure.
- II) La 'determination' devient determined.
- III) La 'durativity' devient non-durative. Cela signifie que l'on se place dans une dimension logique et plus temporelle.

7.3.3.3 Repetition

La repetition est une opération qui sauve (ou prévient) un échec de la fusion perceptive dans le cas **d'ordres de grandeur incompatibles**.

(137) *Marie mangeait à la cantine l'année dernière.*

Dans (137) il y a un problème d'ordre de grandeur entre les images de *manger à la cantine* et de *l'année dernière*. La première prendra environ une heure (plusieurs exceptionnellement) tandis que la seconde s'étale sur des milliers d'heures. L'interprétation d'une occurrence unique de *manger à la cantine* n'est pas possible.

L'interprétation faite est que *Marie* a mangé **plusieurs fois** à la cantine. Au moins deux fois, peut-être même chaque jour de la semaine. La quantité ou le placement des répétitions de *manger à la cantine* n'a pas besoin d'être précisé pour que l'énoncé soit considéré comme correct, tant qu'il y a plusieurs occurrences.

Comment cette opération de répétition est-elle représentée dans le modèle ?

- 1) Un conflit d'ordre de grandeur doit exister.
- 2) L'opération de répétition ne s'applique que sur le AIS dominant, jamais le complément.
- 3) L'ordre de grandeur de l'image identifié dans le AIS dominant doit être inférieur à celui de l'image identifié dans le AIS complément. Par soucis de modularité nous considérons que le module aspectuel demande la permission au module perceptif s'il peut réaliser une répétition, et donc faire une comparaison des ordres de grandeurs (nous voulons éviter d'avoir les informations d'ordre de grandeur dans les deux modules).

Une conséquence de cette interrogation est que l'on ne peut répéter que les AIS qui disposent d'un 'imageID' instancié.

- 4) Le 'viewpoint' est instanciable en une figure.
- 5) La 'multiplicity' est instanciable en une multiple.
- 6) La 'durativity' est instanciable en durative.

Si ces conditions sont remplies, il est possible d'appliquer une opération de répétition, qui consiste à :

- I) Instancier le 'viewpoint' de l'AIS en ground.
- II) Transformer la 'determination' de l'AIS en une variable non instanciée, quelle que soit sa valeur précédant la fusion.
- III) Instancier la 'multiplicity' de l'AIS en multiple.

IV) Demander au module perceptif de mettre à jour 'l'imageID'. La repetition peut transformer l'image en une nouvelle image.

Voici la description d'une fusion qui se sert de l'opération de repetition. L'énoncé traité est le (137) *Marie mangeait à la cantine l'année dernière*. Nous nous plaçons au moment de fusionner *l'année dernière* et *manger à la cantine* (le verbe manger est à l'infinitif).

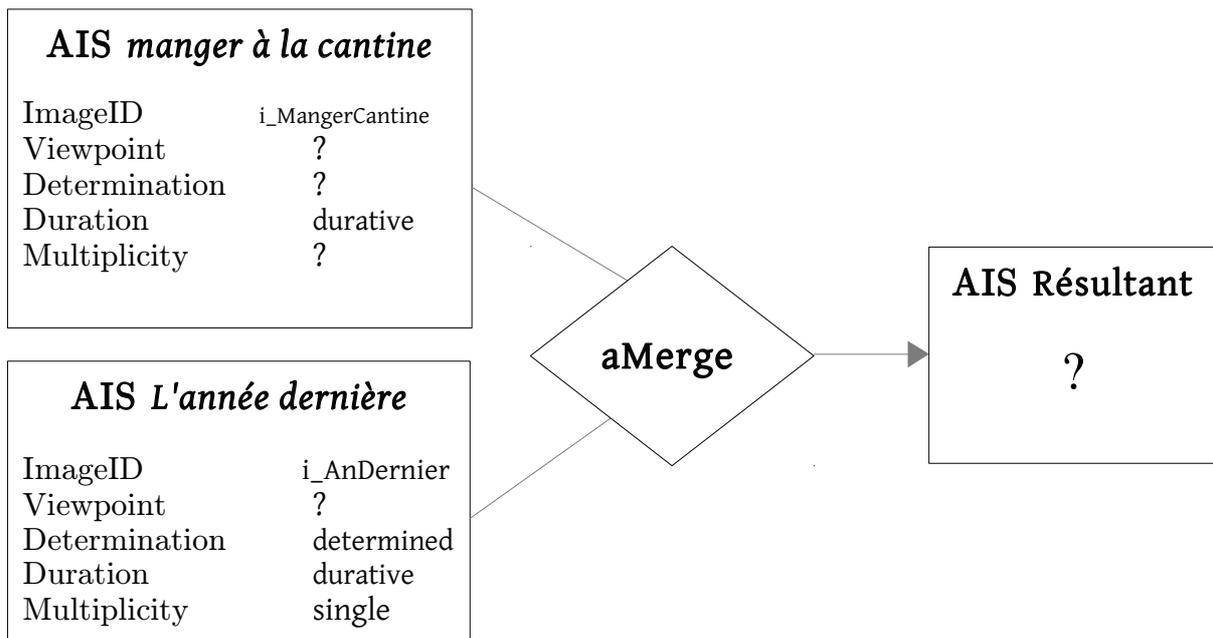


Figure 41 : Avant la fusion des AIS de *manger à la cantine* et *l'année dernière*

Les étapes dans le cas d'une implémentation où la repetition est une opération de récupération sont les suivantes :

- a) Lors de l'unification des 'imageIDs', le module perceptif renvoie un échec en indiquant un conflit d'ordre de grandeur.
- b) Le AIS dominant est *manger_à_la_cantine* [], et son image est bien celle qui a l'ordre de grandeur le plus petit (demande et confirmation du module perceptif).
- c) Le 'viewpoint' de *manger_à_la_cantine* [] est non instancié, il est donc compatible avec une figure.
- d) La 'multiplicity' de *manger_à_la_cantine* [] n'est pas single.

À ce moment, toutes les conditions pour l'application d'une repetition sont valides. On obtient

alors comme nouveau AIS dominant, avant la Fusion :

	imageID	viewpoint	determination	durativity	multiplicity
<i>manger à la cantine</i>	i_MangerCantine	ground	?	durative	multiple

Celui-ci peut alors s'unifier avec l'AIS *L'an dernier*[]. On remarquera que si cette dernière AIS avait disposé d'un 'viewpoint' de type figure, l'unification aurait échoué à cause d'un conflit entre figure et ground. C'est le cas pour l'énoncé suivant :

(138) * *Marie mangeait à la cantine en deux ans.*

La répétition résoudrait le conflit d'ordre de grandeur, mais pas l'unification qui la suit.

7.3.3.4 Slice

Considérons un autre énoncé où un conflit d'ordre de grandeur se manifeste :

(139) *Marie a mangé à la cantine l'année dernière.*

L'interprétation de la répétition existe, mais il existe une seconde interprétation ici qui n'est pas possible avec (137) : Celle où Marie n'a mangé qu'une seule fois à la cantine.

Comment résoudre le conflit d'ordre de grandeur en ayant cette interprétation ? Et pourquoi l'interprétation d'occurrence unique n'est-elle pas possible avec (137) ? Répondons d'abord à la première interrogation.

On considère que si Marie n'a mangé qu'une seule fois, ce n'est pas pendant toute la durée d'une année, mais uniquement pendant une petite partie de celle-ci, d'une durée équivalente à son repas. Dans cette interprétation, c'est *l'an dernier* qui est "découpé" et qui est considéré comme une partie seulement plutôt que le tout.

Cette sélection d'une partie est appelée le slice. Quand s'applique-t-elle et quelles sont ses effets ?

- 1) Un conflit d'ordre de grandeur doit exister.
- 2) L'opération de slice s'applique sur le AIS avec un identifiant vers l'image ayant le plus grand ordre de grandeur. Le slice peut s'appliquer sur le AIS dominant ou le complément (contrairement à la répétition qui n'est applicable qu'au dominant).

L'opération commence (comme la répétition) en demandant au module perceptif la permission de

"slicer" l'image. Cette requête implique aussi que la perception soit capable de distinguer l'image (pour permettre le point 3).

- 3) Le 'viewpoint' est instanciable en un ground.
- 4) La 'determination' est instanciable en determined.
- 5) La 'durativity' est instanciable en durative.

Si ces conditions sont remplies, il est possible d'appliquer une opération de slice, qui consiste à :

- I) Instancier le 'viewpoint' de l'AIS en figure.
- II) Transformer la 'determination' de l'AIS en une variable non instanciée, quelle que soit sa valeur précédent le slice.
- III) Demander au module perceptif de mettre à jour 'l'imageID'. Le slice transforme l'image en une partie de l'image.

Voici la description d'une fusion qui se sert de l'opération de slice. L'énoncé traité est le (139) *Marie a mangé à la cantine l'année dernière*. Nous nous plaçons au moment de fusionner *L'année dernière* et *manger à la cantine* (le verbe manger est à l'infinitif) (voir Figure 41 : Avant la fusion des AIS de manger à la cantine et l'année dernière).

Les étapes dans le cas d'une implémentation où le slice est une opération de récupération sont les suivantes :

- a) Lors de l'unification des 'imageIDs', le module perceptif renvoie un échec en indiquant un conflit d'ordre de grandeur.
- b) l'AIS dont l'image a le plus petit ordre de grandeur est *L'an_dernier[]*. Le module perceptif confirme l'ordre de grandeur ET la possibilité de "slicer" l'image.
- c) Le 'viewpoint' de *L'an_dernier[]* est non instancié, il est donc compatible avec un ground.
- d) La 'determination' de *L'an_dernier[]* est bien determined.

À ce moment, toutes les conditions pour l'application d'un slice sont valides. On obtient alors comme nouveau AIS dominant, avant la Fusion :

	imageID	viewpoint	determination	durativity	multiplicity
<i>L'année dernière</i>	i_AnDernier	figure	?	durative	single

Celui-ci peut alors s'unifier avec le AIS *manger_à_la_cantine[]*.

Nous nous étions demandé pourquoi l'interprétation d'occurrence unique, avec un slice donc, n'est pas possible avec (137).

(137) *Marie mangeait à la cantine l'année dernière.*

Les AIS fusionnés pour décrire repetition et slice sont exactement les mêmes pourtant ! En fait, l'échec pour (137) arrive lors de la fusion suivante, lorsque l'on veut fusionner avec l'AIS ci-dessous :

	imageID	viewpoint	determination	durativity	multiplicity
Imparfait	?	ground	?	?	?

Un slice de *l'année dernière* conduirait à avoir un AIS résultant avec un 'viewpoint' de valeur figure. Unifier figure et ground conduit à un échec et il n'existe pas d'opération qui puisse récupérer le traitement. Heureusement, puisque l'interprétation d'une occurrence unique n'est pas valide : Il serait possible de faire un focus-out, mais on ne parviendrait pas à terminer le traitement de l'ensemble de la phrase sans lancer une seconde prédication. L'AIS finale aurait un ground, ce qui pose problème si l'on n'est pas dans une narration ou un dialogue⁴³.

7.3.3.5 Shift

Le Shift est la cinquième opération identifiée du aMerge. Celle-ci consiste à faire glisser une situation dans le temps, c'est-à-dire la déplacer d'une certaine durée par rapport à une "référence". Voyons quelques exemples pour bien identifier le phénomène :

(140) *Jean gravira le sommet en trois jours.*

(141) *Jean atteindra le sommet en trois jours.*

(142) *Jean gravira le sommet dans trois jours.*

(143) *Jean atteindra le sommet dans trois jours.*

Dans (140) la situation *gravir le sommet* **durera** trois jours, dans (141) la situation *atteindre le sommet* ne **se réalisera qu'au bout de** trois jours. (141) a une interprétation équivalente à (143). Attention cependant, une différence d'utilisation entre (141) et (143) existe : dans le premier

⁴³ La nécessité d'avoir une figure dans l'AIS finale de l'énoncé traité n'est pertinente que si l'on reste dans le cadre de la phrase. Dans ce cas une AIS finale avec un ground sera perçue comme incomplète.

énoncé le "présent" est amovible (On peut commencer la phrase par "*L'année prochaine*"), alors que le second ne peut se dire que si on reste "ancré" au présent d'énonciation (On ne peut pas commencer la phrase par "*L'année prochaine*").

La correspondance que nous mentionnons entre (141) et (143) n'existe pas entre (140) et (142), puisque (142) peut se comprendre comme *Jean commencera son ascension dans trois jours* ou *Jean finira son ascension dans trois jours*. Il n'y a pas d'information sur la durée que Jean prendra pour *gravir le sommet*.

Pour résumer, seuls les énoncés (141), (142) et (143) déplacent la réalisation d'une partie ou de la totalité de la situation d'une durée de trois jours. On parle de **shift**. La durée dans (140) décrit le temps de l'exécution de la situation, pas le positionnement de son exécution.

Le shift semble systématique avec la préposition *dans*, mais pas avec *en*. On peut penser que la télicité de la situation joue un rôle dans les différences d'interprétations. Mais ce n'est pas tout :

(144) *Jean atteindra le sommet dans les trois jours.*

(144) N'a pas le même sens que (143), et en fait les *trois jours* est une période pendant laquelle la situation aura lieu. Il n'y a pas de shift.

(145) **Jean atteindra le sommet dans le voyage.*

Si avec (144) la préposition *dans* se comportait comme un *pendant*, dans (145) ce n'est plus possible. L'énoncé n'est pas correct.

L'interprétation d'énoncés avec les prépositions *en* et *pendant* (entre autre, on peut aussi citer *il y a*, *depuis*, *jusque*, etc.) demande de prendre en compte de nombreux éléments : Le 'viewpoint' (contraint par la télicité) et la 'Détermination', ainsi que des concepts liés à la temporalité : "l'ancrage", "le présent d'énonciation" et même la "direction" du shift (afin de pouvoir faire la différence entre *il y a* et *dans* ; entre *depuis* et *jusque*. On peut ainsi percevoir le shift comme une paire d'opérations).

Nous avons identifié certaines des contraintes d'applications du shift, ainsi que certains de ses effets. En ce qui concerne les informations relatives à la temporalité que nous n'avons pas développées dans ce chapitre, nous sommes malheureusement obligé de les renvoyer dans les perspectives futures...

La formalisation du shift est donc à prendre comme une partie seulement de l'opération

complète.

Voici les contraintes d'applications sur l'AIS du complément :

- 1) Le 'viewpoint' est instanciable en une figure.
- 2) La 'determination' est instanciable en undetermined (En pratique cela signifie que ce sont essentiellement les "durées", c'est-à-dire les situations non distinguables, qui seront la cible d'un shift).
- 3) La 'durativity' est instanciable en durative.

Si ces conditions sont remplies, il est possible d'appliquer à l'AIS du complément une opération de shift qui consiste à :

- I) Transformer le 'Viewpoint' de l'AIS en une variable non instanciée, Quelle que soit sa valeur précédant la fusion.
- II) Transformer la 'determination' de l'AIS en determined.
- III) Transformer la 'durativity' de l'AIS en non-durative.
- IV) Transformer l'image_id existante en la fusionnant avec une image_id de **i_after**.

Voyons en pratique comment l'opération s'exécute sur l'énoncé (141), au moment de la fusion entre *en trois jours* et *atteindre le sommet* :

(141) *Jean atteindra le sommet en trois jours.*

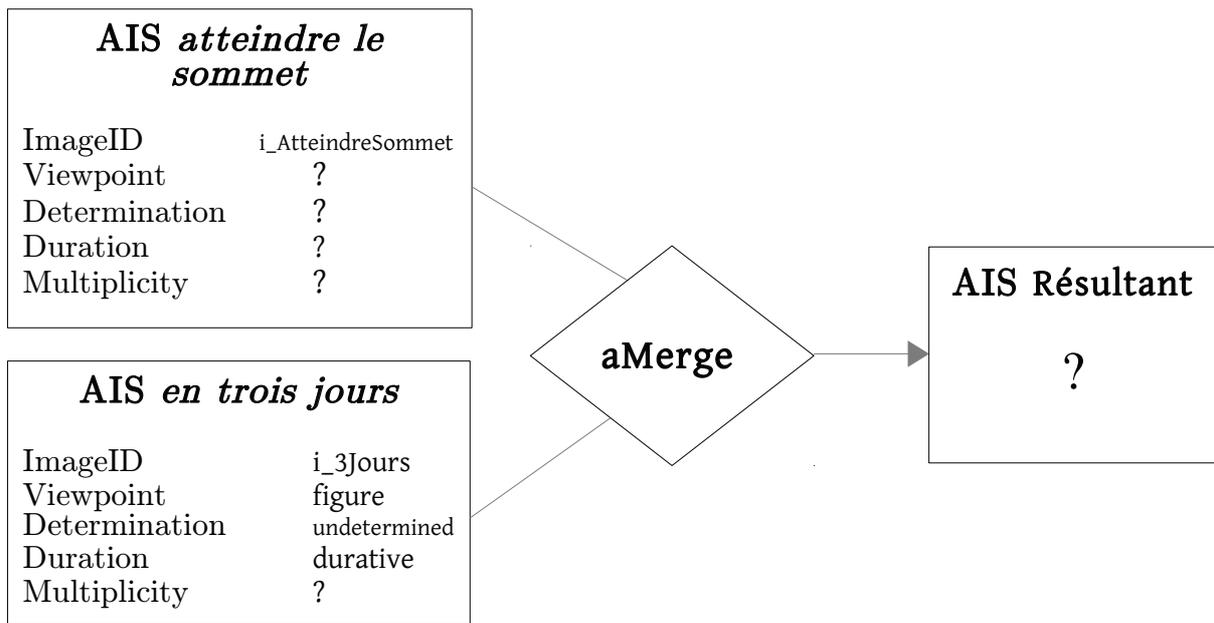


Figure 42 : Avant la fusion des AIS de *atteindre le sommet* et *en trois jours*

- a) Si on ignore les 'imageIDs', *atteindre le sommet* peut s'unifier avec n'importe quelle AIS a priori. Le module perceptif va placer une contrainte de figure sur le 'viewpoint', mais cela ne pose pas de problèmes dans l'immédiat.
- b) C'est un conflit d'ordre de grandeur qui empêche la fusion de *i_3jours* et *i_AtteindreSommet*. Aucun slice ne peut s'appliquer ici, et la seule répétition envisageable (*atteindre le sommet* plusieurs fois !) ne permet pas d'unifier ensuite avec l'AIS de *en trois jours*.
- c) Une possible unification ne fonctionne que si un focus-out s'applique sur l'AIS de *atteindre le sommet* et un shift sur l'AIS de *en trois jours* :

	imageID	viewpoint	determination	durativity	multiplicity
<i>atteindre le sommet</i>	i_Atteindre Sommet	?	?	?	?

Devient après un focus-out :

	imageID	viewpoint	determination	durativity	multiplicity
<i>atteindre le sommet</i>	i_Atteindre Sommet	figure	determined	non-durative	?

Tandis que

	imageID	viewpoint	determination	durativity	multiplicity
<i>en trois jours</i>	i_3jours	figure	undetermined	durative	?

Devient après un shift :

	imageID	viewpoint	determination	durativity	multiplicity
<i>en trois jours</i>	i_after_3jours	?	determined	non-durative	?

L'unification entre les deux AIS après opérations devient possible. Aucun conflit d'ordre de grandeur n'apparaît en raison de la 'durativity' non-durative.

Attention, au moins deux informations sont sous-enseignées dans notre modèle sur les effets d'un shift :

- Nous ne représentons pas l'information de "positionnement" (postérieur et ultérieur) de l'image résultante par rapport à l'image "d'origine". Cela signifie que ne différencions pas *en* ou *dans* de *il y a* ; *ni depuis* de *jusque* ; *avant* de *après* ; etc.
- Nous pensons que le "positionnement" et l'image "d'origine" (nous avons parlé aussi "d'ancre" lors de nos recherches) font partie d'une AIS **augmentée** intégrant le traitement de la temporalité !

On notera que chez Gosselin l'énoncé serait résolu par un "déplacement" de l'intervalle de l'événement *atteindre le sommet* ; tandis que notre opération de shift s'applique sur *en trois jours* : c'est l'unification des deux AIS qui permet à l'événement de se retrouver déplacé dans le futur.

7.4 Lecture des arbres de AIS

Le modèle construit un arbre de AIS en réalisant un aMerge à chaque nœud. Cet arbre sera plus ou moins grand selon le nombre et la taille des énoncés, et la capacité de mémorisation du système : Nous considérons que tout système, qu'il soit humain ou informatique va détruire les AIS et stocker (dans une mémoire longue durée) une partie des informations de l'arbre. Nous ne connaissons pas et ne proposerons pas de règles pour diriger la destruction des AIS "obsolètes", mais insistons sur le fait qu'il est possible de le faire à tout moment. Il est d'ailleurs tout à fait envisageable d'avoir des destructions systématiques dès qu'un aMerge est terminé et de reconstruire l'arbre à la volée si c'est nécessaire (lorsque l'on pose une question d'interprétation précise à un humain, il peut prendre quelques secondes pour répondre).

Mais plaçons-nous avant ces destructions. Les aMerges sont des opérations et ne sont donc pas stockées (à l'exception de l'exécution de l'opération focus-out, mais nous renvoyons son stockage à un autre module). Nous avons un arbre de AIS avec une AIS dit "finale", pour le traitement d'un énoncé. Comment lire et interpréter ces informations ?

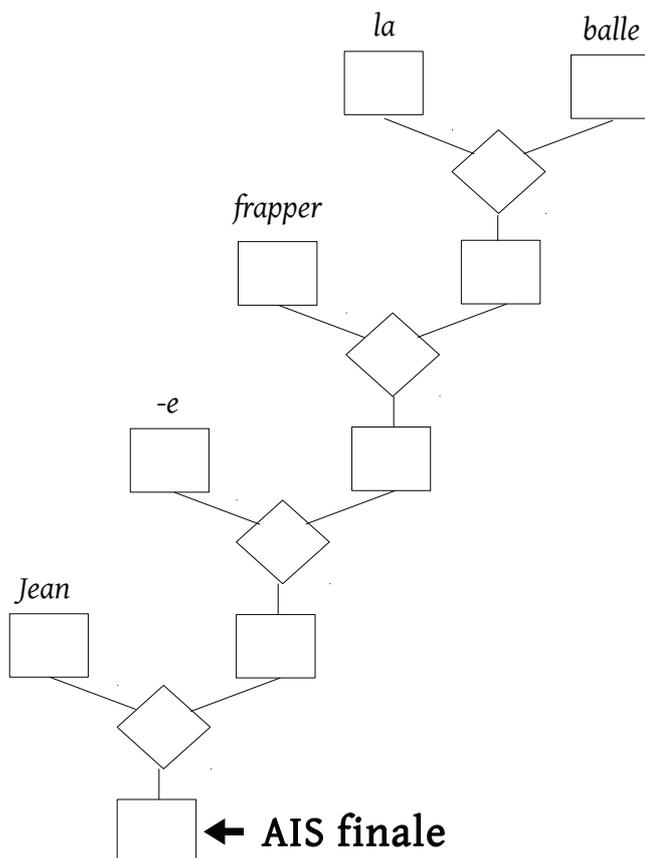


Figure 43 : AIS finale d'un énoncé

L'AIS "finale" ne fournit pas toutes les informations que l'on souhaiterait avoir sur un énoncé. L'information de répétition par exemple, que l'on obtient grâce à la 'multiplicity', n'est accessible que dans l'AIS où l'opération de répétition s'est appliquée. Il faut donc disposer d'une partie précise de l'arbre pour pouvoir connaître les détails des interprétations d'un énoncé.

Lorsque nous avons l'AIS, comment – en tant qu'observateur externe - lire ces résultats ? Par exemple, comment savoir qu'une situation est perçue comme achevée ou pas ? L'AIS fournit la réponse grâce au 'viewpoint' : une figure signifie que les bornes sont perçues, et donc que la situation est perçue comme achevée. Un ground signifie que les bornes ne sont pas connues, et

donc que la situation n'est probablement pas achevée. Voici une liste des points d'interprétations possibles d'un énoncé :

- L'énoncé est **correct** si tous les éléments fournis par la syntaxe ont pu être fusionnés avec succès et si l'AIS finale a une figure comme 'viewpoint'. Par exemple : *Marie a couru le marathon en trois heures.*
- L'énoncé est correct mais incomplet si tous les éléments fournis par la syntaxe ont pu être fusionnés avec succès mais que l'AIS finale a un ground comme 'viewpoint'. Par exemple : *Marie courait dans le parc pendant une heure [, puis allait ...].* En général ces énoncés deviennent corrects si on les fait précéder d'une apposition, ou si on les place dans une narration.
- La situation (*courir dans le parc* ou *courir un marathon*) est perçue comme **achevée** si le 'viewpoint' de sa AIS (c'est-à-dire l'AIS qui correspond à l'intitulé de la situation) est une figure ; ou si au contraire elle peut toujours avoir lieu ou est **inachevée**, maintenant ou dans le futur (sans certitudes) : elle a alors un ground comme 'viewpoint'.
- La situation est **répétée** si la 'multiplicity' de sa AIS est à multiple. Sinon la situation est **unique**. Il est tout à fait possible d'avoir plusieurs répétitions qui porteront sur des AIS différentes, comme dans *Marie mangeait du gâteau en été.* (En considérant des étés génériques.
- La période (*été, trois minutes, etc.*, on fait ici référence au complément circonstanciel de temps) est **entière**, perçue dans son ensemble (durée exacte), s'il n'y a pas eu de slice de l'image perceptive. L'information est stockée dans le module perceptif mais peut-être retrouvée dynamiquement par le module aspectuel en résolvant le aMerge concerné. La période est **partielle** (une partie d'un été peut être un jour) s'il y a eu un slice.
- On peut savoir si la période correspond à une durée ou une période "générique" (*n'importe quel été*) si la 'determination' de sa AIS est undetermined, on dira qu'elle est **indéterminée** ; ou si au contraire la période est connue est identifiable parmi d'autres si la 'determination' de sa AIS est determined, elle est déterminée.

8 Traitements

8.1 Exemples de traitements

Nous venons de voir les attributs et opérations du modèle au cas par cas. Il est temps de décrire son comportement sur un énoncé complet et plus seulement sur de petits bouts de phrases choisies.

8.1.1 Un cas détaillé

L'énoncé que nous allons traiter est celui-ci :

(146) *Marie a mangé à la cantine en 2012.*

Le temps de conjugaison est le passé composé. L'action de manger à la cantine peut avoir eu lieu une seule fois pendant l'année 2012, ou de manière répétée sans que l'on sache exactement à quelle fréquence (régulièrement ou seulement deux ou trois fois) de l'année.

On notera qu'il n'est pas possible de remplacer *2012* par *un an* : on obtiendrait un énoncé incorrect.

Voici un arbre syntaxique simplifié qui guidera l'ordre des fusions successives. La simplification que nous avons réalisée porte sur *manger à la cantine* que l'on traitera directement comme un syntagme (puisque les aMerges qui s'y rapportent ne comportent que des unifications de variables non-instanciées et de 'imageID' que nous ne détaillons pas).

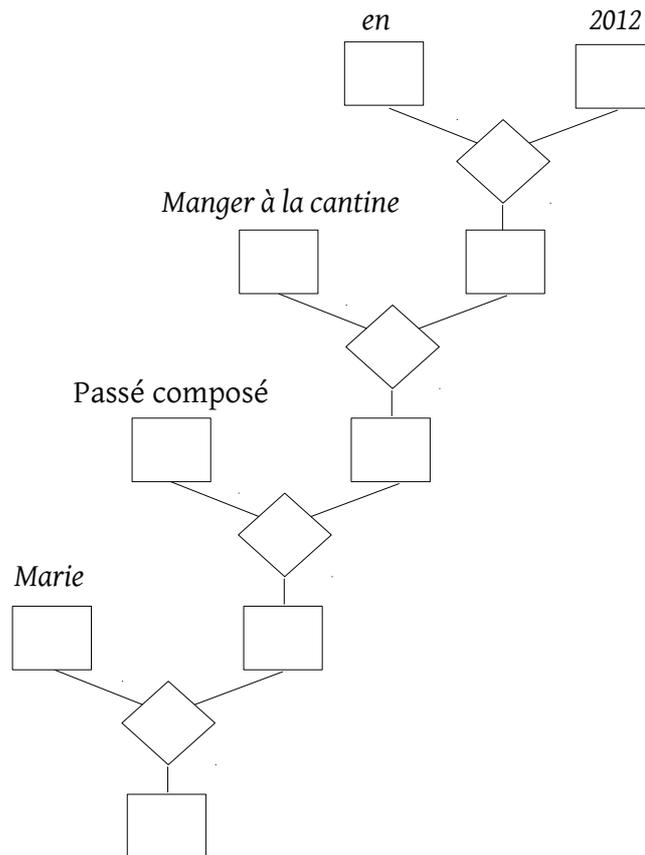


Figure 44 : Arbre syntaxique de *Marie a mangé à la cantine en 2012*

En 2012

La première fusion syntaxique prend en entrée les unités lexicales *en* et *2012*. Puisque la fusion syntaxique dirige la fusion sémantique, notre module aspectuel recevra les données suivantes :

en (dominant), ***2012*** (complément)

Le module aspectuel va trouver les AIS correspondantes de chacune des unités. Voici les AIS de *en* et *2012* :

	imageID	viewpoint	determination	durativity	multiplicity
<i>en</i>	i_simul	figure	?	?	?
<i>2012</i>	i_2012	?	determined	?	?

Le **aMerge** s'exécute avec ces deux AIS. On va tout d'abord tenter d'**unifier** (page 133) les deux AIS :

i_simul + i_2012

Pour unifier i_simul et i_2012, le module aspectuel les transmet au module perceptif, qui lui renvoie un 'imageID' correspondant. La fusion perceptive (page 98) n'est pas décrite par le modèle, aussi nous nous contenterons de supposer que le nouvel 'imageID' sera **i_2012**.

figure + ?

L'unification d'une figure avec une variable non instanciée réussit. La variable non-instanciée et l'unificateur (le résultat de l'unification) sont des figures. On a : **figure + figure → figure**.

? + determined

Tout comme l'unification précédente, l'unification de la 'determination' réussit. On a **determined + determined → determined**.

? + ?

Aucunes des 'durativity' et 'multiplicity' ne sont instanciées. Le module se contentera de garder un unificateur non-instancié.

L'AIS résultante de l'unification est celle-ci :

	imageID	viewpoint	determination	durativity	multiplicity
<i>en 2012</i>	i_2012	figure	determined	?	?

Le module la garde en mémoire jusqu'à ce qu'on l'utilise pour le prochain aMerge

Manger à la cantine en 2012

La prochaine paire syntaxique identifiée est celle-ci :

manger à la cantine (dominant), ***en 2012*** (complément)

Les AIS correspondantes :

	imageID	viewpoint	determination	durativity	multiplicity
<i>manger à la cantine</i> (0)	i_mangerCantine	?	?	?	?

en 2012 (0)

imageID	viewpoint	determination	durativity	multiplicity
i_2012	figure	determined	?	?

En tant que verbe avec complément, *manger à la cantine* a une AIS avec uniquement son 'imageID' instancié. Sa télicité (voir page 104) n'est pas connue à ce moment.

Le chiffre entre parenthèse (ici un zéro) après la partie de l'énoncé indique une version particulière de l'AIS au cours du traitement (avant ou après fusion ; ou si on a plusieurs AIS possibles).

Le **aMerge** s'applique, avec la phase d'unification :

i_mangerCantine + i_2012 → error

En l'état, le module reçoit une erreur du module perceptif, provoquée par l'incompatibilité des ordres de grandeurs des images correspondantes (voir page 101).

Dans notre implémentation, les opérations du aMerge (voir page 136) s'appliqueront toutes sans avoir été "appelées", certaines réussiront à lever l'erreur et d'autres non. Pour simplifier notre exemple, nous contenterons de décrire les opérations nécessaires pour la résolution du conflit d'ordre de grandeur.

Focus-in et Slice

On réalise tout d'abord un focus-in (page 137) sur l'AIS *en 2012(0)*[], possible puisque l'on a les valeurs figure et determined et que l'imageID' est instancié. Comme résultat le 'viewpoint' devient un ground.

On réalise ensuite un slice de l'AIS obtenue, ce qui possible puisque l'on a les valeurs ground, determined et qu'il est possible d'instancier en durative. l'AIS obtenue est la suivante :

en 2012 (1)

imageID	viewpoint	determination	durativity	multiplicity
i_part_of_2012	figure	?	?	?

Le changement le plus important est réalisé lors de la mise à jour de 'l'imageID' (par le module perceptif) : On ne parle plus de la totalité de l'année 2012 mais d'une partie.

L'unification entre les AIS *manger à la cantine*[] et *en 2012(1)*[] est possible sans provoquer de conflit d'ordre de grandeur : En effet une partie de 2012 peut très bien s'accommoder de la durée d'un repas à la cantine. Voici l'AIS résultante dans ce cas :

manger à la cantine
en 2012 (1)

imageID	viewpoint	determination	durativity	multiplicity
i_mangerCantine_part_of_2012	figure	?	?	?

On a perdu la valeur *determined*, ce qui signifie que l'on peut très bien parler de *manger à la cantine* comme une situation indéterminée.

Repetition et focus-in

Il existe une deuxième solution pour résoudre le problème d'ordre de grandeur : appliquer l'opération de répétition (page 139). Celle-ci est possible parce que l'AIS *manger à la cantine(2)* est dominante, que le module perceptif est capable de construire une image de *manger à la cantine plusieurs fois*, et que l'on peut instancier *figure*, *multiplicity* et *durative*. Voici l'AIS résultante⁴⁴ :

	imageID	viewpoint	determination	durativity	multiplicity
<i>manger à la cantine (2)</i>	i_mangerCantine_repeated	ground	?	?	multiple

Le conflit d'ordre de grandeur est résolu, mais un conflit d'unification est apparu : le *ground* de *manger à la cantine (2)* est incompatible avec la figure de *en 2012 (0)*. L'opération de *focus-in* intervient alors. Celle-ci nous permet de retrouver une AIS avec un *ground* :

	imageID	viewpoint	determination	durativity	multiplicity
<i>en 2012 (2)</i>	i_2012	ground	determined	?	?

Nous pouvons unifier les deux AIS et obtenir une seconde interprétation. On notera que *multiple* n'est pas transmis à l'AIS résultante comme expliqué page 136 :

	imageID	viewpoint	determination	durativity	multiplicity
<i>manger à la cantine en 2012 (2)</i>	i_mangerCantine_repeated_2012	ground	determined	?	?

L'unification des 'imageID' se fait correctement.

i_mangerCantine + i_2012 → i_mangerCantine_part_of_2012

i_mangerCantine + i_2012 → i_mangerCantine_repeated_2012

Il existe une troisième interprétation possible où nous avons fait un *focus-in* et un *slice en plus* de la répétition. Sa AIS est celle-ci, bien que nous ne la développerons pas dans les aMerges suivants (son comportement n'est pas différent des autres) :

	imageID	viewpoint	determination	durativity	multiplicity
<i>manger à la cantine en 2012 (2)</i>	i_mangerCantine_repeated_part_of2012	ground	determined	?	?

⁴⁴ Voir page 126 si la redondance de l'information de répétition (une dans l'image, une dans l'AIS) vous dérange.

A mangé à la cantine en 2012

Nous arrivons au troisième aMerge. Voici les entrées :

Passé composé (dominant), *manger à la cantine en 2012* (complément)

l'AIS du temps de conjugaison :

	imageID	viewpoint	determination	durativity	multiplicity
passé composé fr.	?	figure	?	?	?

Il y a deux AIS possibles pour *manger à la cantine en 2012*.

	imageID	viewpoint	determination	durativity	multiplicity
<i>manger à la cantine en 2012</i> (1)	i_mangerCantine_part_of_2012	figure	?	?	?

OU

	imageID	viewpoint	determination	durativity	multiplicity
<i>manger à la cantine en 2012</i> (2)	i_mangerCantine_repeated_2012	ground	determined	?	?

Lors du aMerge, l'unification se passe sans problèmes entre les AIS *passé composé fr.* et *manger à la cantine en 2012* (1). On obtient :

	imageID	viewpoint	determination	durativity	multiplicity
<i>a mangé à la cantine en 2012</i> (3)	i_mangerCantine_part_of_2012	figure	?	?	?

figure + ground → error

Il y a par contre un conflit entre les 'viewpoints' lors de l'unification entre *passé composé fr.* et *manger à la cantine en 2012* (2). Est-ce que cela signifie que cette seconde interprétation est disqualifiée ? Pas si une opération peut "sauver" l'unification, en l'occurrence l'opération de focus-out (page 138) :

focus-out

Un focus-out peut s'appliquer sur *manger à la cantine en 2012* (2) puisque 'l'imageID' est instanciée, ground et durative sont instanciables, et enfin qu'aucun focus-out n'a été utilisé auparavant dans la phrase. l'AIS résultante est la suivante :

	imageID	viewpoint	determination	durativity	multiplicity
<i>manger à la cantine en 2012</i> (4)	i_mangerCantine_repeated_2012	figure	determined	non-durative	?

L'unification avec l'AIS du passé composé peut avoir lieu, le conflit d'unification entre figure et ground étant levé.

La valeur non-durative que l'on retrouve ici indique que la situation est forcément prise d'un point de vue "logique", afin de pouvoir tirer des conclusions du type : "Elle a mangé à la cantine en 2012 et n'aurait pas dû" ou "Elle a mangé à la cantine en 2012 plutôt que chez lui". On obtient comme AIS résultante :

	imageID	viewpoint	determination	durativity	multiplicity
<i>a mangé à la cantine en 2012 (4)</i>	i_mangerCantine_repeated_2012	figure	determined	non-durative	?

On notera qu'il n'est pas possible d'appliquer de focus-in sur l'AIS du passé simple, puisque l'imageID' n'est pas instanciée.

Marie a mangé à la cantine en 2012

Le dernier aMerge, avec l'identification de l'AIS de *Marie*[] est une unification simple, sans conflits, puisque l'AIS de Marie ne contient qu'un 'imageID' sans ordre de grandeur conflictuel. Aussi nous allons reprendre les trois AIS finales :

	imageID	viewpoint	determination	durativity	multiplicity
<i>Marie a mangé à la cantine en 2012 (A)</i>	i_mangerCantine_part_of_2012	figure	?	?	?
	imageID	viewpoint	determination	durativity	multiplicity
<i>Marie a mangé à la cantine en 2012 (B)</i>	i_mangerCantine_repeated_2012	figure	determined	non-durative	?
	imageID	viewpoint	determination	durativity	multiplicity
<i>Marie a mangé à la cantine en 2012 (C)</i>	i_mangerCantine_repeated_part_if_2012	figure	determined	non-durative	?

Ce sont les trois interprétations possibles de l'énoncé *Marie a mangé à la cantine en 2012* :

- Marie a mangé à la cantine une seule fois à un moment de l'année 2012.
- Marie a mangé à la cantine à plusieurs reprises pendant toute l'année 2012.
- Marie a mangé à la cantine à plusieurs reprises pendant une partie de l'année 2012.

8.1.2 Exemple commenté

(147) *Il parlera en 10 minutes*

L'exemple (147) a deux interprétations possibles :

- Soit *parler* correspond à une sorte de discours ou de présentation, et dans ce cas la durée de sa prise de parole sera typiquement de 10 minutes.
- Soit *parler* correspond à *avouer*, lors d'un interrogatoire, et dans ce cas *parler* commencera dans 10 minutes.

L'image de *parler* est différente selon le contexte, mais dans les deux cas le module perceptif répondra 'oui' au moment de l'unification des 'viewpoints' à la question sur la télicité de l'image. La seconde interprétation ne peut s'obtenir que si un 'focus-out' s'applique sur l'AIS de *parler* : en effet la prédication rend l'AIS non-durative et permet alors une application d'un shift sur l'AIS de *en 10 minutes*.

8.2 Implémentation

8.2.1 Précisions techniques

Le modèle a fait l'objet d'une implémentation. Les deux langages informatiques utilisés sont prolog et python.

La Figure 45 est un schéma de l'architecture globale du programme. Nous signalons les correspondances avec les modules théoriques (que l'on pourra retrouver page 88).

Le programme est utilisable au travers d'une interface HTML (gérée avec du code python et le mini-serveur python "bottle"). Bien que largement améliorable⁴⁵, l'interface autorise une visualisation des résultats du traitement sous la forme de tableaux. La Figure 46 est un exemple de la fusion entre les énoncés *manger à la cantine* et *en 2010*.

⁴⁵ Dans l'idéal, les résultats d'interprétations finaux devraient proposer des paraphrases (mode texte) et un graphique de type "ligne du temps" ; deux lectures beaucoup plus aisés pour le néophyte.

Module lexico-syntaxique

tmt_lexic.pl
Gestion des entités lexicales.
tmt_syntax.pl
Gestion des règles syntaxiques et de leur exécution (parsage tabulaire Ascendant : Algorithme CYK).
lex_file.pl
Base des données lexicales.
syn_file.pl
Base des données (règles) syntaxiques.

tmt_syntax_to_semantic.pl
Interface entre la syntaxe et le module temporelle (détection de la correspondance entre entités lexicales et AIS)

Module aspectuel

tmt_model.pl
Algorithme principal : aMerge.
tmt_semantic.pl
Gestion des unités sémantiques.
tmt_ss.pl
Gestion des AIS (unités sémantiques spécialisées).
sem_file.pl
Base des données sémantiques symboliques (liste des AIS connues).

tmt_semantic_to_perception.pl
Interface entre le module de sémantique symbolique et le module perceptif.

Module perceptif

tmt_perception.pl
Gestion des informations Perceptives. Fusion perceptive.
per_file.pl
Base de données des informations perceptives.

tmt_spy.pl / tmt_store.pl / tmt_util.pl
Fichiers utilitaires (log, sauvegarde, debug)

Couche Python
tmt.py / tmt_util.py

tmt_prolog.py
Gestion du code prolog (du modèle) par soumission de prédicats et récupérations de structures (logs, AIS).
tmt_interface.py
Interface graphique. Création dynamique de pages HTML.

Figure 45 : Schéma de l'implémentation

Nous donnons en annexe des extraits de certains des fichiers.

Veillez entrer une phrase :

en2	+ d2010	=>	i1 i 2010	
[,f ,u ,]	+ [i 2010 , ,d ,]			
[,f ,d ,]	+ [i 2010 ,f ,d ,]	=>	[i 2010 ,f ,d ,]	← Le "en" fu
en	+ d2010	=>	i2 i 2010	
[, ,u ,]	+ [i 2010 , ,d ,]			
[, ,d ,]	+ [i 2010 , ,d ,]	=>	[i 2010 , ,d ,]	← Le "en" _u
i1 i 2010	+ manger a la cantine	=>	i3 i 2010 i manger_cantine1	
[i 2010 ,f ,d ,]	+ [i manger_cantine , , ,]			
[i 2010 ,f ,slice ,d ,]	+ [i manger_cantine ,f ,d ,]	=>	[i 2010 i manger_cantine1 ,f ,d ,]	← Le resultat du merge entre "en 2010" (fu) et "manger a la cantine"
i2 i 2010	+ manger a la cantine	=>	i4 i 2010 i manger_cantine1	
[i 2010 , ,d ,]	+ [i manger_cantine , , ,]			
[i 2010 ,f ,slice ,d ,]	+ [i manger_cantine ,f ,d ,]	=>	[i 2010 i manger_cantine1 ,f ,d ,]	← Les deux resultats possibles du merge entre "en 2010" (_u) et "manger a la
i2 i 2010	+ manger a la cantine	=>	i5 i 2010 i manger_cantine1	
[i 2010 , ,d ,]	+ [i manger_cantine , , ,]			
[i 2010 ,g ,d ,]	+ [i manger_cantine ,g ,multiple ,d ,]	=>	[i 2010 i manger_cantine1 ,g ,d ,]	

Figure 46 : Interface

8.2.2 Résultats

L'implémentation nous a essentiellement permis de tester le modèle sur quelques exemples choisis. Le module syntaxique (voir en annexe [syn_file.pl](#), page 194) et la base de données sémantique (voir en annexe [per_file.pl](#) et [sem_file.pl](#)), bien que réduites, nous permettent déjà de traiter correctement (c'est-à-dire d'obtenir les mêmes interprétations qu'un locuteur humain francophone) la combinatoire des énoncés constructibles. Nous fournissons un échantillon des interprétations réalisées dans l'annexe [14.1 Listes d'exemples et de leurs interprétations](#), page 188.

Nous considérons que l'accroissement des données lexicales, syntaxiques et perceptives⁴⁶ est réalisable sans obstacles majeurs, permettant ainsi le traitement davantage d'énoncés.

Un travail plus conséquent et plus à même de contraindre le modèle à évoluer, est le renseignement des AIS des entités grammaticales et lexèmes de la "classe fermée". Si nous avons testé certains mots comme *avant*, *après*, *en train de*, *quand*, *toujours*, etc., ceux-ci n'ont pas bénéficié de l'exhaustivité des tests de *en* et *pendant*. L'une des explications est l'importance de la temporalité dans leur traitement. Pour cela, le modèle ne permet pas en l'état de prendre en compte les distinctions temporelles. Il est par exemple capable de traiter l'imparfait et le passé composé comme des AIS renseignant le viewpoint, mais en ignorant totalement les informations de passé, d'ancrage dans le présent, de révolu, etc.

46 En tout cas si on se limite à un module symbolique simplifié de la perception !

9 Calcul de la parcimonie

Nous avons insisté tout au long de ce document sur l'objectif de parcimonie d'un modèle de traitement du temps (et de tout modèle cognitif en général). Maintenant que nous avons présenté notre modèle, comment pouvons-nous **mesurer** sa parcimonie ?

Nous proposons plus bas une liste d'énoncés en français et allons calculer le nombre de bits nécessaires pour le traitement de leurs aspects par deux modèles : Le null-model (voir chapitre [2.1 La parcimonie](#)) et le nôtre.

Le null-model est le modèle extrême de non-parcimonie, tandis que notre modèle en est une tentative parcimonieuse : C'est la comparaison des bits nécessaires de chaque modèle qui permet de dire si un modèle est parcimonieux ou pas. Nous nous intéressons à la compression du modèle par rapport au null-model, c'est-à-dire à la compression au sens de la complexité de Kolmogorov, et non au temps d'exécution ! Nous nous limitons ici à l'aspectualité, mais il est entendu que la démarche de calculer la parcimonie peut être réalisée pour tous les modèles computationnels, qu'ils traitent de temporalité en plus, de modalité, ou du langage en général.

L'objectif du calcul de la parcimonie n'est pas de montrer que nous faisons mieux que le null-model, ce qui n'est guère difficile ; L'idéal aurait été de se comparer aux autres modèles, mais malheureusement ceux-ci ne proposent pas pour le moment une telle mesure. Notre objectif dans ce chapitre est de proposer pour la première fois un calcul de la parcimonie d'un modèle du traitement de l'aspect, afin de mettre en avant la notion même de parcimonie qui nous a servi à à formuler des critiques à l'encontre de différents modèles (chapitre [5 Des modélisations pour le traitement de l'aspect](#)).

9.1 Liste d'énoncés

Les énoncés peuvent avoir l'une des deux formes suivantes :

[sujet][situation][tense][préposition][déterminant B][circonstanciel]

Ou en ayant le circonstanciel de temps en apposition :

[circonstanciel], [sujet][situation][tense][préposition]

Par exemple :

(148) *Marie a mangé du gâteau pendant ces trois heures.*

Ou

(149) *Cet été, Marie a couru un marathon.*

Tous les énoncés sont construits en combinant les valeurs possibles de chaque élément, en se limitant à quelques valeurs différentes listées dans le tableau ci-dessous et à une possible absence de l'élément. Les énoncés grammaticalement incorrects sont disqualifiés.

La valeur d'apposition signifie que nous plaçons le complément circonstanciel en tête avant une virgule, comme dans (149). Voici les différentes valeurs utilisées.

Tableau 6 : Composants des exemples

Sujet	Situation	Tense	Prép.	Dét.	Circonstanciel	Apposition ?
Marie	Manger du gâteau	imparfait	∅	∅	été	non
	Courir un marathon	Passé composé	en	ces/cet	trois mois	oui
			pendant		trois heures	
					trois minutes/ secondes	
					hier	

En prenant la première valeur de chaque élément, le premier énoncé correct que nous obtenons est :

Marie mangeait du gâteau en été.

En changeant "l'action" on obtient :

Marie courait un marathon en été.

Un énoncé (grammaticalement) incorrect est :

* *Marie mangeait du gâteau pendant été.*

Ces listes d'éléments et la structure des énoncés nous permettent d'obtenir **52 énoncés différents et grammaticalement corrects** (on considère que le module syntaxique a déjà défaussé les énoncés grammaticalement faux ; Ceux-ci ne rentreront donc pas dans nos calculs) dont nous donnons un extrait plus bas (la liste complète se trouve en annexe page 188). On peut statuer pour chaque énoncé sur plusieurs points (voir chapitre [7.4 Lecture des arbres de AIS](#) page 148 pour des explications supplémentaires) :

- Si l'énoncé est **correct** (pas seulement syntaxiquement mais surtout "sémantiquement").
- Si "l'action" est unique ou **répétée**.
- Si l'ensemble de la situation est unique ou **répétée**.
- Si la période (*été, trois minutes, etc.*) est perçue dans son ensemble (durée exacte) ou partiellement (une partie d'un été peut être un jour).
- Si la période correspond à une durée ou une période "générique" (*n'importe quel été*), ou un période connue bien précise (voir chapitre [7.2.5 Détermination temporelle](#), page 123).
- Sur quelle partie de l'énoncé porte la prédication. On parlera ici d'**attitude**.

Il est possible d'avoir plusieurs interprétations pour un énoncé, chaque ligne correspondant alors à une interprétation.

La liste des énoncés se structure de cette manière :

Énoncé	Répétition "action"	Répétition de l'ensemble de la situation	Ensemble de la période	Période déterminée	Attitude	Commentaires
--------	---------------------	--	------------------------	--------------------	----------	--------------

La partie "commentaires" contiendra d'autres remarques parfois spécifiques à l'énoncé, interprétables à partir des AIS et aMerge réalisés par le modèle. Même si nous ne faisons pas apparaître ces "points" dans une colonne, ceux-ci seront pris en compte dans le calcul de la complexité. Il s'agit de :

- Déterminant quantificateur de l'objet de l'action (par exemple dans *Marie a couru un*

marathon en trois mois, "un" doit être quantitatif).

- "Prédication" obligatoire, ce qui correspond à des interprétations d'interdits, d'exceptions, ou d'échos à un énoncé précédent.
- Si le contexte est suffisant ou pas. S'il est nécessaire de connaître l'ordre de grandeur de l'AIS finale précédente (comme une apposition par exemple : *À cette époque*) ou si l'énoncé est le début d'une liste d'actions dans une narration.
- Aspect terminatif ou inchoatif⁴⁷.

Bien évidemment, nous sommes bien conscient que la liste sur laquelle nous évaluons ici le modèle n'est qu'un petit échantillon des énoncés possibles d'une langue. Pour rappel, la liste des énoncés se trouve page 188 dans l'annexe 14.1 Listes d'exemples et de leurs interprétations.

9.2 Complexité du null-model

Nous allons tout d'abord calculer la complexité du null-model pour être capable de fournir les interprétations aspectuelles correctes (que nous avons proposées dans le chapitre précédent) des 52 énoncés. Rappelons que le null-model est le cas extrême où le modèle est inexistant, et où on se contente de lier la liste des énoncés possibles à la liste des interprétations possibles (apprentissage par cœur).

Nous avons distingué six types d'éléments syntaxiques (voir Tableau 6) en entrée, ce qui correspond à un simple raccourci puisque les aMerges non-pertinents sont déjà réalisés. Cette simplification n'impacte pas le calcul de la parcimonie d'un modèle du traitement du temps et sert uniquement à éviter de surcharger les exemples.

Si on a **N exemples** en entrée, et **K interprétations** possibles, le null-model nécessite $S=N \times K$ bits : il faut K bits pour indiquer si chaque interprétation est valide ou non. Cette valeur croît linéairement avec N.

⁴⁷ Inchoatif indique que le commencement de la situation est déplacé vers le futur, par exemple : *Il va courir dans cinq minutes*.

Dans notre cas, **N a une valeur de 52.**

K est calculé en multipliant le nombre de solutions possibles pour chacun des "traits" d'interprétations listés dans la partie précédente (tous les traits n'étant pas binaires, cela complique un peu la tâche), ce qui nous donne :

$$1+(2 \times 2 \times 2 \times 2 \times 2 \times 2 \times 6 \times 3) = 1153 = \mathbf{K}$$

La première valeur 1 correspond à la valeur incorrecte (si incorrectes, les valeurs suivantes ne sont pas utiles).

Les 2 correspondent aux traits binaires.

6 correspond à l'absence de prédication plus les 5 prédications possibles identifiées pour ces énoncés, à savoir "l'action", la période, la période+déterminant, la situation sans le temps grammatical, la totalité de l'énoncé (par exemple dans *Marie a mangé du gâteau cet été*, la prédication, ou attitude, porte obligatoirement sur *manger du gâteau*. Sauf si la phrase est une réponse ou un écho : "- *Marie n'a pas mangé du gâteau depuis janvier.* - *Ah si ! Elle a mangé du gâteau **cet été**.*)

3 correspond à l'interprétation complète ou l'une des deux incomplètes : C'est-à-dire soit qu'il manque une information contextuelle, comme dans l'exemple (150). Soit que l'on attend une énumération, ou une liste, comme dans l'exemple (151).

(150) *Elle mangeait à la cantine.* (quand cela ? *L'année dernière*)

(151) *Elle allait à la montagne en hiver.* (et à la plage en été)

Pour le null-model, avec notre listes d'exemples, il faut $52 \times 1153 = 59\ 956$ bits uniquement pour relier les quelques énoncés à leurs possibles interprétations aspectuelles.

Si **les exemples peuvent être décrits avec P bits**, et qu'il peut exister un modèle qui prend des décisions correctes à partir de P marqueurs binaires, alors un null-model de ce type nécessite $S = 2^P \times K$ bits plus les bits nécessaires pour décrire le lexique. $2^P \times K$ ne croît pas avec N. Dans le cadre de nos exemples, l'approche n'est parcimonieuse que si P est petit (puisque N est lui aussi déjà petit). Il serait donc intéressant de trouver P marqueurs binaires pertinents, mais uniquement si P reste petit ! Si P augmente, en augmentant le nombre de traits par exemple, l'intérêt du modèle diminue en même que le taux compression se réduit !

Pour calculer P, on additionne le nombre de traits possibles pour chaque type d'élément : traits

des noms, des verbes (dynamacité, ponctualité, télélicité, genres, ...), des temps, des prépositions, etc.

$$\mathbf{P} = P_{\text{noms}} + P_{\text{verbes}} + P_{\text{tense}} + P_{\text{prep}} + \dots$$

L'idéal est bien entendu de limiter au maximum la taille de P. Cette réduction est plus intéressante que le null-model sans traits : En effet, si nous n'avons que peu d'énoncés pour notre démonstration, en pratique le nombre d'énoncés N va tendre vers l'infini...

L'enjeu est ensuite de diminuer **la taille S (en bits) de l'algorithme** qui produit les bonnes interprétations à partir des P marqueurs binaires. La taille S est celle de la procédure qui permet de passer des exemples aux interprétations. Celle-ci doit introduire une réduction de complexité en étant inférieure à une simple mise pair à pair ($S \leq 2^P \times K$).

Une version moins stricte associe une **combinaison de traits** à chaque interprétation, toutes les autres combinaisons étant considérées comme incorrectes. L'idée est de se demander à partir de chaque interprétation "quels traits" déclenchent cette interprétation.

Ce modèle statique (au sens où il ne crée pas d'autres résultats qui ne soient déjà fournis en entrée) a besoin (en plus de S) de $K \times P$ bits en entrée. Si une interprétation peut être obtenue à partir de différents énoncés, il faut la dupliquer, ce qui augmente K.

Si on introduit un masque qui peut utiliser la valeur "non-définie" en plus des deux valeurs "définies", on trouve qu'un tel modèle nécessiterait $S = \log_2(3^P) \times K$ bits.

C'est une approche purement structuraliste (appariement de structure, sans calculs dynamiques) qui n'a (pour le moment) pas réussi à expliquer les phénomènes temporels du langage. Bien que sa complexité peut être réduite (si P reste petit), nous doutons qu'un tel modèle existe en conservant à K une valeur raisonnable.

9.3 Complexité du modèle avec AIS et aMerge

9.3.1 Les forces du modèle

Les cas qui viennent d'être présentés sont des cas extrêmes de complexité et ne représentent que

la limite haute de complexité pour un modèle. Tous les modèles scientifiques sont censés être plus parcimonieux que le null-model. Notre motivation est de trouver un modèle plus parcimonieux parmi tous les autres (Nous sommes conscient qu'être parcimonieux sur l'aspect ne signifie pas qu'on le sera sur la temporalité ou la modalité).

Le calcul de la complexité d'un modèle n'étant pas chose courante, nous ne pourrions nous comparer qu'au null-model, mais espérons que ce souci de parcimonie sera partagé avec les modèles existants et futurs sur le temps.

L'idée de l'AIS consiste à réduire les P marqueurs en une structure de p bits pour chaque élément. Considérons donc **p la taille de l'AIS**. Une procédure de type **aMerge** produit son résultat considérant seulement 2p bits : Elle prend en entrée uniquement deux AIS. La taille de cette structure est intéressante d'un point de vue de la complexité dès que $2p < P$. On a donc $S < 2^{2p} \times K$. Si une procédure aMerge est possible, le modèle gagne en parcimonie. La taille S du mécanisme d'unification déclenché par le aMerge, s'il ne demande pas d'exécuter des opérations, est considérée comme négligeable (algorithme récursif d'unification).

Attention cependant, les modèles à base de marqueurs doivent assigner ces marqueurs aux éléments du lexique. Ainsi la complexité du modèle est l'addition de la complexité et l'algorithme et de la complexité du lexique (c'est-à-dire des marqueurs associés au lexique) :

Complexité = complexité de l'algorithme + complexité du lexique.

$$C = S + L$$

À l'échelle d'un langage complet, le nombre d'éléments est très important⁴⁸. Fort heureusement pour nous, les lemmes qui fournissent une information temporelle précise sont beaucoup plus rares : Ils sont membres de ce que l'on appelle "la classe fermée" (voir page 25). Ceux-ci disposent chacun d'une AIS complète. Tous ces autres éléments disposent chacun d'une AIS où **seule l'imageID est renseignée**, ce qui permet d'économiser la taille d'une AIS à chaque fois, mais pas les "informations perceptuelles" comme l'ordre de grandeur et la télicité. Notons que ces deux informations perceptuelles concernent essentiellement les syntagmes verbaux (*Manger une pomme a un ordre de grandeur temporel, pomme n'en a pas*).

48 D'après l'Académie française (http://academie-francaise.fr/la-langue-francaise/questions-de-langue#56_strong-em-nombre-de-mots-de-la-langue-franaise-em-strong) le "français fondamental" comprend entre 1000 et 3000 entrées, tandis que les dictionnaires de la langue courante comportent environ 60 000 entrées.

On imagine aisément que cela représente une très grande quantité d'information, mais sans meilleure explication, tous les modèles existants en sont (plus ou moins) dépendants ! Dans notre cas, ces informations sont renvoyés dans le module perceptif ; module qui contiendrait déjà une quantité d'informations bien plus riche encore à propos du lexique.

Mais quelle est la taille p d'une AIS ?

9.3.2 Occupation des AIS et informations perceptuelles

Une AIS est constituée de quatre switches ('viewpoint', 'determination', 'durativity', 'multiplicity'), qui prendront chacun moins de 2 bits (pour les deux valeurs plus la non-instanciation), et un "pointeur" ('imageID').

L'occupation du pointeur est difficile à estimer, aussi nous préférons représenter dans notre calcul les attributs effectivement nécessaires au module perceptif plutôt que de passer par le pointeur. Ces attributs sont la télécité d'une image (implémentée comme un switch - voir chapitre [6.3.2 L'image perceptive](#) - mais le module perceptif ce contente de répondre de manière binaire (oui/non) aux interrogations du module temporel) qui tiendra sur 1 bit, et l'ordre de grandeur (l'âge estimé de l'univers a un ordre de grandeur de 17,6 dans notre implémentation, mais on considérera pour la démonstration que l'humain n'est pas capable de percevoir la différence entre 1 milliard et 15 milliards d'années, et nous voyons large !) qui tiendra sur 4 bits.

En résumé :

4 switches ('viewpoint', 'determination', 'durativity', 'multiplicity') qui prennent $4 \times 1,5 = 6$ bits

1 réponse binaire (la télécité) qui prend 1 bit

1 ordre de grandeur qui prend 4 bits

Ce qui donne pour une AIS et ses informations perceptuelles $8+1+4 = 11$ bits

9.3.3 Calcul de la complexité de l'algorithme

L'enjeu final est la taille de l'algorithme de aMerge. L'usage de l'unification est extrêmement économe, tandis que les exceptions coûtent cher.

Si on a **R exceptions statiques**, alors il faut au plus $S=2^p \times R \times p$ bits (2^p puisque l'on utilise deux

AIS pour chaque fusion, ou aMerge, et p puisque l'on obtient une AIS comme résultat) pour repérer la configuration qui déclenche chaque exception (ou $S = \log_2(3^{2p}) \times R \times \log_2(3^p)$ si l'on autorise des masques). L'emploi de masques permet de diminuer la valeur de R.

Si l'on néglige qu'il existe une procédure d'unification (qui est une simple copie), on a $S \leq 2p \times R \times p$. On peut imaginer une procédure plus compacte qu'une simple table pour produire les exceptions à partir des $2p$ bits.

Dans la pratique, ces R exceptions correspondent chez nous aux AIS qui peuvent déclencher des **opérations** (chapitre [7.3.3 Opérations](#), page 136), et sont en nombres équivalents (si on se sert de masques) : 5.

En résultat, la complexité maximale S de notre algorithme est de :

$$S = \log_2(3^{2 \times 11}) \times 5 \times \log_2(3^{11}) = 35 \times 5 \times 18 = 3150 \text{ bits}$$

9.4 Conclusion sur la parcimonie

Ces valeurs de la complexité sont comme des benchmarks du modèle. Elles permettent de comparer son efficacité en complément de la comparaison traditionnelle du nombre de résultats corrects (de bonnes interprétations).

Nous faisons mieux que le null-model, ce qui doit être le cas de n'importe quelle modèle ! Mais nous ne nous sommes pas comparés à la complexité d'autres modèles : Ceux-ci (ceux présentés dans le chapitre [5 Des modélisations pour le traitement de l'aspect](#), par exemple) ne prennent pas (pour le moment) la complexité de leur algorithme en considération.

Cela ne veut pas dire qu'ils ne sont pas parcimonieux ! Mais sans calculs, et sans "bornes" clairement définies, il y a toujours le risque d'observer un accroissement incontrôlé de la complexité ; c'est typiquement le cas des modèles qui attachent des algorithmes aux éléments du lexique (de Saussure, 2003; Gosselin, 1996; Schilder, 2004).

Nous invitons dans ce chapitre à prendre la parcimonie d'un modèle (aspectuel ou non) en considération afin de pouvoir évaluer les modèles entre eux autrement que par leurs résultats.

Nous pensons que la prise en compte de la parcimonie amènera à contraindre utilement l'élaboration et l'amélioration des modèles du traitement du langage. L'espoir est de voir les modèles se rapprocher d'un traitement cognitif et au final de donner de meilleurs résultats.

10 Limites et Perspectives

Nous arrivons à la fin de la présentation de notre modèle, et il est temps de revenir sur les résultats obtenus, ses points forts et ses limites.

10.1 Production et interprétation

Notre système global pour le traitement du langage a été présenté comme devant être capable de réaliser l'interprétation et la production d'énoncés (sous réserve bien entendu de disposer des informations perceptuelles utiles). Un tel système utiliserait un seul module par "domaine", par exemple la temporalité ou juste l'aspect, et non pas deux distincts qui traiteraient l'un l'interprétation et l'autre la production !

Malgré nos intentions, l'état actuel du modèle et son implémentation ne sont pas capables de produire des énoncés. Ce n'est que faute de temps et nous sommes confiant pour pouvoir adapter le modèle dans cette optique.

10.2 la temporalité et l'aspect

La plus grande lacune de notre modèle, nous l'avons exprimé dès le début du document, est de ne proposer qu'un traitement de l'**aspect**, et pas de **temporalité** (et encore moins de la **modalité**).

Si on se suffit au cadre étroit d'une situation qui n'utilise que des simultanités entre périodes/événements, que l'on oublie que la temporalité ou des notions comme la référence de Reichenbach existent, alors notre modèle est suffisant.

La mise en pratique montre que ce cadre limité peut parfaitement servir à mettre en avant une

démarche, mais qu'il est insuffisant à exprimer toute l'étendue de l'aspectualité : temporalité, aspect et modalité sont étroitement imbriqués ! Si l'étude de *en* et *pendant* (par exemple) fonctionne jusqu'à un certain point, on parvient aux limites de notre cadre lorsque l'on aborde l'inchoativité et la terminativité. Des prépositions tels que *depuis*, *jusqu'à*, *avant*, *après*, impliquent des positionnements par rapports à des situations. C'est aussi le cas des temps grammaticaux, qu'ils soient simples ou composés. Alors nous pouvons le représenter en nous servant de 'l'imageID' et de la puissance du module perceptif, mais sans avoir pu s'en **servir** pendant la fusion.

Il se trouve que, au-delà des opérateurs unaires relationnels ('imageID'), nous avons travaillé sur l'intégration de la temporalité dans notre modèle. L' AIS c'est vue enrichie de quelques attributs supplémentaires :

- 'Anchor', qui correspond à l' AIS d'ancre qui permet de distinguer *dans* (l'ancre est le "présent relatif") de *après* (l'ancre peut être redéfinie). Par exemple *Il part dans dix minutes* et *Il part dix minutes après le spectacle*. Nous sommes tentés de comparer cette 'Anchor' au temps de référence de Reichenbach.

- 'Gone' (révolu) et 'Distance' afin d'exprimer le positionnement par rapport au présent d'énonciation, comme nous l'avons suggéré dans le chapitre [4.2.1 Passé-Présent-Futur](#) page 55.

Avec l'ajout de ces trois attributs, nous sommes alors capable de traiter et distinguer des situations courantes comme *hier/la veille/le jour d'avant*. Toutes ces situations peuvent représenter le même jour, mais elles ont un comportement différent.

Si le temps nous a malheureusement manqué pour formaliser ces pistes, il nous a été suffisant pour entrevoir les possibilités et le potentiel de notre approche, l'utilisation d'une structure sémantique (AIS, qui serait alors augmentée en TIS, "T" pour Temporal, ou ATIS) et d'une fusion sémantique (aMerge), tout en nous appuyant sur les travaux et idées de nos prédécesseurs.

Des travaux de Roser Saurí (Saurí & Pustejovsky, 2012) ont montrés qu'il était possible d'exprimer la factualité (factuality) avec une combinaison de seulement trois variables binaires (des switches ?), avec la même idée de parcimonie que nous avons suivie.

Nous pensons que le plein potentiel du modèle, et sa validation, passent obligatoirement par

l'intégration de la temporalité et de la modalité.

10.3 Le module perceptif

Le second obstacle majeur pour permettre au modèle de passer à l'échelle (pour une utilisation pratique dans le Traitement Automatique du Langage) est le flou du module perceptif (chapitre [6.3 Le module perceptif](#)). Cette "boîte noire" va sans aucun doute demander beaucoup plus de travail que la temporalité, puisqu'elle fait appel aux connaissances générales du monde.

Une première approche pour formaliser le module perceptif serait de se servir d'ontologies construites "à la main". Des informations d'homogénéité, d'ordres de grandeur (relatives, et pas absolues comme nous le faisons), mais aussi certainement de dynamique, de causalité, etc. devront y être introduites. Est-il possible d'utiliser les ontologies existantes et d'en extraire les informations voulues ? Les résultats seraient-ils "complets" ? Dans notre cas, le minimum serait de pouvoir répondre aux interrogations du module aspectuel (temporel) : être capable de comparer les ordres de grandeurs de deux "images", savoir si une "image" donnée est téléique (non homogène, revoir chapitre [4.1.6 Téléicité](#), page 48), pouvoir fournir une nouvelle "image" à partir d'une paire fournie.

La seconde approche serait d'ancrer dans le monde un "agent", comme le sont les humains. L'agent, un robot par exemple (voir entre autres Kaplan, 2001 pour l'acquisition du langage chez des robots) acquerrait ses connaissances lors de ses interactions avec l'environnement et construirait sa "base de donnée", son propre module perceptif. Le travail est vaste, peut-être encore plus que pour les ontologies, mais on aurait là un système en évolution permanente beaucoup plus proche du fonctionnement humain. Comme pour les ontologies, la formalisation d'un modèle sémantique comme le nôtre fournit de précieuses indications pour les interactions entre modules (comparaisons des ordres de grandeurs, connaissance sur l'homogénéité d'une situation).

Dans les deux cas, il est impératif de connaître les interactions des modules sémantiques (c'est ce que nous avons fait dans le chapitre [7.3.1 Interactions du module aspectuel](#), page 131) avec le

"module perceptif" pour le contraindre et savoir ce qu'il doit être capable de faire ; et à l'inverse pour réviser les modules sémantiques, si nécessaire.

10.4 Les dimensions sémantiques

Aspect et temporalité sont des catégories d'une "dimension" temporelle. Cette dimension est une manière de percevoir un objet ou une situation : l'*été* est ainsi une période de temps d'une durée de trois mois.

Mais il est possible de passer par d'autres "dimensions" (Gärdenfors, 2000, 2014), et de devoir réaliser des calculs très différents. C'est ce que nous avons montré avec la prédication (chapitre 7.3.3.2 Focus-out et "prédication" temporelle, page 138) qui "fait perdre" la dimension temporelle en met en avant une **autre** dimension. La prédication de *été* nous conduira par exemple à percevoir la situation non pas comme une période de trois mois, mais comme une situation "qui n'est pas l'hiver, l'automne, ou le printemps". On perd la "durativité" (chapitre [7.2.6 Durativité](#), page 125) pour obtenir une information différente.

On se retrouve ainsi, au cours de notre calcul de l'aspect, à devoir prendre en compte des dimensions non-temporelles. On se retrouve avec des attributs de l'AIS qui eux aussi peuvent être partagés avec d'autres modules (dimensions) : la 'multiplicity' (les nombres et quantités), le 'viewpoint' (directement inspiré de la perception et de la dimension spatiale) et la 'determination'.

Ces imbrications et partages d'attributs entre modules semblent être un pas vers une plus grande parcimonie du système cognitif en général, et même des validations pour la pertinence d'utiliser ces attributs : En effet, si différents modules qui n'ont a priori pas beaucoup en commun (le module de l'aspect et la quantification des choses par exemple) utilisaient la même information (l'homogénéité), l'existence de celle-ci en serait renforcée. On veut souligner ici que sa valeur est nécessaire aux calculs, et n'en est pas le résultat.

11 Conclusion

11.1 Apports de la thèse

L'objectif de cette thèse a été de mettre au point et d'implémenter un modèle **parcimonieux** et **cognitivement plausible** du traitement de l'aspect dit "**grammatical**" dans les langues naturelles. Nous nous sommes concentré sur le français – bien que gardant un œil sur l'anglais et même quelques langues qui nous sont moins familières (roumain, russe).

- Nous avons souhaité remettre en avant l'idée de **plausibilité cognitive** dans le traitement du langage, en faisant intervenir quelques-uns des arguments existants (dont certains sont issus d'autres domaines que la linguistique): "pauvreté du stimulus", approche évolutionnaire, refus d'un langage de la pensée (language of thought), modularité des capacités cognitives.
- Nous avons défini et formalisé la contrainte de **parcimonie** en proposant que celle-ci devienne une contrainte forte dans la modélisation du traitement du langage. Chacun des choix d'implémentation ont été faits pour être le plus parcimonieux possible : Les résultats obtenus montrent qu'il est tout à fait possible de l'être en traitant un grand nombre de situations.
- En concordance avec la volonté de plausibilité cognitive et de parcimonie, nous intégrons le module aspectuel dans une **architecture générale** du traitement du langage qui prend en compte non seulement la syntaxe mais surtout le module perceptif. Ce module, surtout inspiré du modèle des espaces conceptuels de Gärdenfors, est présenté comme une alternative aux systèmes de connaissances de type ontologiques (système de prédicats logiques, language of thought). Nous détaillons les interactions minimalistes entre le module aspectuel et les autres modules (syntaxique et perceptif).
- L'**Aspectual Informations Structure** (AIS) est le seul et unique type de structure du

module aspectuel. Celle-ci répond au critère de parcimonie en étant de taille fixe et sans la possibilité de se contenir elle-même. Sa taille reste relativement petite puisqu'elle ne contient qu'un identifiant et quatre **switch**, un switch étant une variable binaire qui peut changer de valeur en cours de traitement et prendre un état "non-instancié" supplémentaire.

- L'**Aspectual Merge** (aMerge, ou fusion aspectuelle) est le seul et unique algorithme contenu dans le module aspectuel. Celui-ci consiste essentiellement en un mécanisme de **fusion** de deux AIS, et à cinq **opérations**. L'aMerge est synchronisé avec la syntaxe et le module perceptif, c'est-à-dire que la fusion aspectuelle (et plus généralement sémantique) est synchronisée avec les fusions syntaxique et perceptive.
- Parmi les opérations de l'aMerge, nous avons découvert un mécanisme intéressant, le focus-out, lié à la **prédication** (ou attitude propositionnelle). Ce phénomène original correspond à la perte de la dimension temporelle d'une partie de la phrase et à l'apparition d'une interprétation "d'exceptionnalité" ou "d'interdit". Nous avons mis en évidence son caractère prédictible, elle est déclenchée selon l'état de structures sémantiques (les AIS), en l'ajoutant au nombre de nos opérations.
- La seconde opération d'importance est le shift, qui est une nouvelle interprétation de ce qui est généralement appelé l'**inchoativité**. Nous montrons comment prévoir ce phénomène en concordance avec l'application du focus-out (ou prédication temporelle).

Dans le cadre de l'aspectualité en français, nous sommes parvenu à réaliser un modèle parcimonieux⁴⁹ capable de prédire l'exactitude ("correct", "incorrect", "bizarre") des énoncés et de fournir toutes les informations utiles pour en déduire les valeurs aspectuelles que sont la répétition, la prise en compte de l'ensemble ou non d'une situation, la perfectivité, la terminativité, l'inchoativité. De plus, l'opération de focus-out a introduit dans le modèle la possibilité d'obtenir des informations précises sur la prédication (ou "l'attitude propositionnelle"), en général non rattachée à l'aspect. Les énoncés que l'on s'était proposé de "résoudre" (page 188) sont tous traités par notre modèle avec les corrections (correct/incorrect) et interprétations attendues.

Ces objectifs ont pu être atteints sans avoir besoin de différencier les "états" (states) et les

49 Sa plausibilité cognitive est encore à débattre.

"événements" (events). Cela a été possible en rejetant le trait de "ponctualité"/"durativité" (que l'on retrouve dans de nombreuses classifications de situations depuis Vendler)⁵⁰.

Si nous nous servons d'ordre de grandeurs, nous évitons de parler de durées et encore moins d'une représentation d'une "ligne temporelle". Nous n'avons pas de "moments", ou "d'intervalles", mais représentons les situations avec une AIS et une "image" du module perceptif. Le seul algorithme est celui du aMerge, aucun algorithme ne se retrouve rattaché à des éléments du lexique.

Avec le aMerge, le modèle peut facilement s'intégrer avec les théories récentes de la syntaxe. L'opération de fusion sémantique (aspectuelle) est parallèle aux fusions syntaxiques. La seule contrainte étant de disposer d'une organisation binaire hiérarchisée des éléments syntaxiques. Cette approche procédurale se distingue par une forte dynamique : Des opérations permettent de modifier les structures en fonction de paramètres simples (valeurs des AIS) ; un mécanisme d'exploration d'arbre permet d'obtenir tous les résultats possibles sur le traitement d'énoncés.

L'état actuel du modèle ne permet pas encore une utilisation opérationnelle en traitement automatique du langage, mais notre objectif a été de montrer qu'il est possible d'obtenir des résultats et des informations aspectuelles pertinentes tout en restant parcimonieux. Ce souci de minimalisme et de plausibilité cognitive nous a certainement permis de dépasser des solutions plus évidentes mais moins parcimonieuses, et ainsi mettre au point un modèle original.

11.2 Avenir

Le modèle demande à connaître les AIS de tous les éléments de la "classe fermée" du temps et de l'aspect pour être exécuté. Cela correspond à moins d'une centaine de mots en français, mais nécessite de nombreux tests sur un maximum d'énoncés pour trouver les valeurs de leurs AIS. La recherche des valeurs sémantiques de tous ces mots va conduire forcément à explorer et intégrer les informations temporelles manquantes : Comment traiter le progressif (anglais) par exemple ? Bien entendu, le travail nécessaire le plus évident est l'étude de la temporalité. C'est bien son intégration dans le modèle qui va déterminer si la piste que nous avons suivie est viable, et si son

50 Tout en gardant néanmoins celui de télélicité ; et sans prendre sur la dynamique.

ajout ne va pas nécessiter des additions incompatibles avec nos contraintes de minimalisme. la temporalité n'est que le début. Nous avons envisagé que nombre d'attributs pourraient être partagés entre les "modules" (l'homogénéité et la 'multiplicity' par exemple), ce qui tendrait à renforcer leur intérêt dans un système minimaliste. Le traitement de la temporalité en plus de l'aspect donnerait ainsi de nombreuses pistes pour les modules suivants à intégrer : modalité, causalité, quantification, homogénéité, espace, etc.

Une application pratique pour le traitement automatique du langage (TAL) ne devient possible que dans ce contexte étendu où (presque) tous les domaines sémantiques sont abordés.

Enfin nous aimerions insister sur un phénomène nouveau et absolument inattendu, ce que nous avons appelé la prédication temporelle : Certainement sous-exploré dans ce document, il est fort probable que ce phénomène ait des liens plus profonds avec le "module perceptif", et soit une composante essentielle d'une modèle comme celui des espaces conceptuels de Gärdenfors.

12 Index

Accomplissements.....	38	Inné.....	23 sv
Achèvements.....	38	Instruction.....	70
Activités.....	39	Intervalle.....	68, 105
Anchoring.....	59, 120	Intervalles.....	55
Apprentissage (enfant).....	23	Kamp.....	74
Borik.....	47	Kolmogorov.....	16 sv
Branching future model.....	34	Le modèle scientifique.....	15
Causalité.....	62	Lemme.....	21
Chaitin.....	16	Ligne temporelle.....	29
Chimpanzé.....	24	Linguistique.....	38
Chomsky.....	23	Linguistique cognitive.....	36
Classes aspectuelles.....	38	Logique formelle.....	29
Classes fermées.....	21	Logique modale.....	29
Classification.....	16	Logique temporelle.....	29
Clauses imbriquées.....	67	Logiques temporelles hybrides.....	33
Complexité.....	16	Machine de Turing.....	18
Comrie.....	40 sv	Metric tense logic.....	31
Conjonction.....	46	Modalité.....	59
Cumulativité.....	49	Modèle statistique.....	23
Declerk.....	48	Modification adverbiale.....	46
Depraetere.....	47	Module perceptif.....	168
Dessalles.....	15	Moens.....	62
Détermination.....	58	Moment d'énonciation.....	52, 54
Différences génétiques.....	24	Moment de référence.....	54, 65
Discourse Representation Structures.....	74	Null-model.....	19
Discourse Representation Theory.....	74	Ockham.....	17
Durativité.....	40	Opérateur binaire.....	114
Dynamicité.....	39, 41	Opérateur unaire.....	108 sv
En.....	71	Opérateurs relationnels.....	56
États.....	38	Ordre de grandeur.....	41
Évolution.....	24	Parcimonie.....	16, 116, 157
Factuality.....	60	Passé.....	52
Figure.....	118	Passé composé.....	9
Filip.....	44	Passé simple.....	9
Focus-in.....	134	Pauvreté du stimulus.....	23
Fréquence.....	50	Pendant.....	71
Futur.....	52	Perfectivité.....	42, 118
Futur périphrastique.....	53	Perspective.....	42
Galton.....	29	Point d'achèvement.....	47
Gène.....	24	Point of event.....	54
Genre.....	20	Point of reference.....	54
Ghadakpour.....	25	Popper.....	16
Gosselin.....	39, 52, 55, 68	Pragmatique.....	62
Granularité.....	25	Précédence.....	105
Ground.....	114	Prédication".....	135
Habitude.....	50	Présent.....	52
Homogénéité.....	33, 49	Prior.....	32
Hwang.....	67	Procès.....	38, 56
Idéaliste (philosophie).....	29	Progressif.....	46
ImageID.....	117	Pustejovski.....	60
Imparfait.....	9	Quine.....	49
Inchoativité.....	46	Réaliste (philosophie).....	29

Réfutabilité.....	16	Système langagier.....	23
Reichenbach.....	35, 52, 54, 64	Talmy.....	44
Reinhart.....	56	Télicité.....	45, 131
Relation.....	106	Temps de conjugaison.....	7
Répétition.....	50	Tense.....	52
Résolution des conflits.....	72	Tense absolu.....	53
Reyle.....	74	Topologie.....	25
Ritchie.....	62	Trait.....	20 sv
Saint Augustin.....	28	Un autre argument sur la plausibilité cognitive est celui de la granularité :.....	25
Saurí.....	60	Unification.....	130
Schubert.....	67	Vendler.....	38, 102
Shift.....	141	Viewpoint.....	114
Smith.....	43	Zénon.....	26
Speech time.....	52, 54	44
Steedman.....	62, 102		

13 Bibliographie

- Allen, J. F. (1984). Towards a general theory of action and time. *Artificial intelligence*, 23(2), 123–154.
- Asher, N. (1993). *Reference to abstract objects in discourse* (Vol. 50). Kluwer Academic Pub.
- Augustin, S. (1993). *Confessions*. Gallimard.
- Auriat, N. (1996). *Les défaillances de la mémoire humaine: aspects cognitifs des enquêtes rétrospectives* (Vol. 136). Ined.
- Bates, E., Bretherton, I., & Snyder, L. S. (1991). *From first words to grammar: Individual differences and dissociable mechanisms* (Vol. 20). Cambridge University Press.
- Borik, O. (2006). *Aspect and reference time*. Oxford University Press.
- Chomsky, N. (1970). Remarks on nominalization. *RA Jacobs, PS Rosenbaum. London, Reading in English Transformational Grammar*, 184–221.
- Chomsky, N. (1988). *Language Problems of Knowledge: The Managua Lectures* (Vol. 16). MIT press.
- Chomsky, N. (1995). *The minimalist program*. Cambridge (ma): MIT press.
- Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems* (Vol. 2). Cambridge University Press.
- Comrie, B. (1985). *Tense*. Cambridge University Press.
- Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, 14(04), 597–650.
- Davidson, D. (1967). The logical form of action sentences. *Essays on actions and events*, 5, 105–148.
- Declerck, R. (1979). Aspect and the bounded/unbounded (telic/atelic) distinction. *Linguistics*, 17(9-10), 761–794.
- Depraetere, I. (1995). On the necessity of distinguishing between (un) boundedness and (a) telicity. *Linguistics and philosophy*, 18(1), 1–19.

- De Saussure, L. (2003). *Temps et pertinence: éléments de pragmatique cognitive du temps*. De Boeck Supérieur.
- Dessalles, J. L. (2008). *La pertinence et ses origines cognitives - Nouvelles théories*. Paris: Hermes-Science Publications.
- De Swart, H. (1998). Aspect shift and coercion. *Natural Language & Linguistic Theory*, 16(2), 347–385.
- Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ* (Vol. 7). Springer.
- Fauconnier, G. (1994). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge University Press.
- Filip, H. (1999). *Aspect, Situation Types and Noun Phrase Semantics*. Ph. D. Thesis, University of California at Berkeley. 1993.[Published as *Aspect, Eventuality Types and Nominal Reference*. 1999. New York/London: Garland Publishing, Inc.].
- Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. The MIT Press.
- Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Clarendon Press/Oxford University Press.
- Fodor, J. A., Garrett, M. F., Walker, E. C., & Parkes, C. H. (1980). Against definitions. *Cognition*, 8(3), 263–367.
- Gärdenfors, P. (2000). *Conceptual Spaces: The Geometry of Thought*. MIT press.
- Gärdenfors, P. (2014). *Geometry of Meaning*. Cambridge, MA; London, England: The MIT Press.
- Gerasymova, K., Spranger, M., & Beuls, K. (2012). A Language Strategy for Aspect: Encoding Aktionsarten through Morphology. In L. Steels (Ed.), *Experiments in Cultural Language Evolution* (Vol. 3, pp. 257–276). Amsterdam: John Benjamins.
- Ghadakpour, L. (2003). *Le système conceptuel, à l'interface entre le langage, le raisonnement et l'espace qualitatif: vers un modèle de représentations éphémères*.
- Gosselin, L. (1996). *Sémantique de la temporalité en français: un modèle calculatoire et cognitif du temps et de l'aspect*. Duculot.
- Gosselin, L. (2005). *Temporalité et modalité*. De Boeck Supérieur.
- Guéron, J. (1993). Sur la syntaxe du temps. *Langue française*, 102–122.

- Hassabis, D., & Maguire, E. A. (2007). Deconstructing episodic memory with construction. *Trends in cognitive sciences*, 11(7), 299–306.
- Hinrichs, E. W. (1985). *A compositional semantics for Aktionsarten and NP reference in English*. Ohio State University.
- Hinrichs, E. W. (1986). Temporal anaphora in discourses of English. *Linguistics and philosophy*, 9(1), 63–82.
- Hurford, J. R. (2003). The neural basis of predicate-argument structure. *Behavioral and Brain Sciences*, 26(3), 261–282.
- Hwang, C. H., & Schubert, L. K. (1992). Tense Trees as the fine structure of discourse. In *Proceedings of the 30th annual meeting on Association for Computational Linguistics* (pp. 232–240).
- Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. Oxford University Press, USA. Retrieved from http://books.google.com/books?hl=en&lr=&id=d9O9-w1c1j4C&oi=fnd&pg=PP5&dq=jackendoff+foundations+of+languages&ots=6mlm5O3DAg&sig=83LhNcMxagy-1c5fjsOCzmEu_5A
- Jackendoff, R. S. (1992). *Semantic structures* (Vol. 18). MIT press.
- Kamp, H. (1968). *Tense logic and the theory of linear order*.
- Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory* (Vol. 42). Dordrecht: Kluwer Academic Publisher.
- Kaplan, F. (2001). *La naissance d'une langue chez les robots*. Paris: Hermes-Science Publications.
- Klein, E. (2004). *Les tactiques de Chronos*. Flammarion.
- Klein, H. G. (1974). *Tempus, Aspekt, Aktionsart*. Niemeyer.
- Klein, W. (1994). *Time in language*. Routledge.
- Kosslyn, S. M. (1996). *Image and brain: The resolution of the imagery debate* (1999th ed.). Cambridge (ma): MIT Press.
- Kowalski, R., & Sergot, M. (1986). A logic-based calculus of events. *New generation computing*, 4(1), 67–95.
- Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event

- semantics. *Semantics and contextual expression*, 75, 115.
- Lakoff, G. (1987). Women, fire, and dangerous things: What categories reveal about the mind.
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford: Stanford university press.
- Leibniz, G. W., & Clarke, S. (2000). *Leibniz and Clarke: Correspondence*. (R. Ariew, Ed.). Hackett Publishing Co.
- Li, M., & Vitanyi, P. M. (1997). *An introduction to Kolmogorov complexity and its applications* (3rd ed.). New York: Springer Verlag.
- Mani, I., Pustejovsky, J., & Gaizauskas, R. (2005). *The language of time: a reader*. Oxford, UK: Oxford University Press.
- Mani, I., & Schiffman, B. (2005). Temporally anchoring and ordering events in news. *Time and Event Recognition in Natural Language*. John Benjamins.
- Mey, J. L. (2001). *Pragmatics: an introduction*. Blackwell publishers Oxford.
- Mikkelsen, T. S., Hillier, L. W., Eichler, E. E., Zody, M. C., Jaffe, D. B., Yang, S.-P., ... Altheide, T. K. (2005). Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*.
- Miyagawa, S., Berwick, R. C., & Okanoya, K. (2013). The emergence of hierarchical structure in human language. *Frontiers in psychology*, 4.
- Moens, M., & Steedman, M. (1988). Temporal ontology and temporal reference. *Computational linguistics*, 14(2), 15–28.
- Mulcahy, N. J., & Call, J. (2006). Apes save tools for future use. *Science*, 312(5776), 1038–1040.
- Munch, D., & Dessalles, J.-L. (2011). Vers un modèle minimaliste du traitement des relations temporelles. In *Actes MFI'11*. Presented at the MFI'11, Rouen.
- Munch, D., & Dessalles, J.-L. (2012). Inferring aspectuality on French sentences: a minimalist approach. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 2055–2060). Retrieved from <http://hal.archives-ouvertes.fr/hal-00723803/>
- Nilsson, N. J. (1998). *Artificial intelligence: a new synthesis*. Morgan & Kaufmann.
- Núñez, R., & Cooperrider, K. (2013). The tangle of space and time in human cognition. *Trends in*

cognitive sciences.

- Parsons, T. (1990). *Events in the Semantics of English*. Cambridge, MA: MIT press.
- Passonneau, R. J. (1988). A computational model of the semantics of tense and aspect. *Computational Linguistics*, 14(2), 44–60.
- Patard, A. (2007). *L'un et le multiple. L'imparfait de l'indicatif en français. Valeur en langue et usages en discours.*
- Prior, A. N. (1967). *Past, present and future*. Oxford: Oxford University Press.
- Prior, A. N. (1968). *Papers on time and tense*. Clarendon Press Oxford.
- Pustejovsky, J., Ingria, B., Sauri, R., Castano, J., Littman, J., Gaizauskas, R., ... Mani, I. (2005). The specification language TimeML. *The language of time: A reader*, 545–557.
- Quine, W. V. (1960). *Word and object* (Vol. 4). MIT press.
- Reichenbach, H. (1947a). *Elements of Symbolic Logic*. New York: The Macmillan Company.
- Reichenbach, H. (1947b). *Elements of Symbolic Logic*. New York: The Macmillan Company.
- Reinhart, T. (1986). States, events and reference time. *Lecture given at the MIT Lexicon project*.
- Reinhart, T. (2000). *Reference time*. Presented at the Language in Use Symposium, Utrecht, Netherland.
- Ritchie, G. D. (1979). Temporal clauses in English. *Theoretical Linguistics*, 6(1-3), 87–116.
- Roberts, W. A. (2007). Mental time travel: Animals anticipate the future. *Current biology*, 17(11), R418–R420.
- Saurí, R., & Pustejovsky, J. (2012). Are you sure that this happened? assessing the factuality degree of events in text. *Computational Linguistics*, 38(2), 261–299.
- Schilder, F. (2004). Extracting meaning from temporal nouns and temporal prepositions. *ACM Transactions on Asian Language Information Processing (TALIP)*, 3(1), 33–50.
- Schilder, F., & Habel, C. (2001). From temporal expressions to temporal information: Semantic tagging of news messages. In *Proceedings of the workshop on Temporal and spatial information processing-Volume 13* (p. 9).
- Slobin, D. I. (1973). Cognitive prerequisites for the development of grammar. *Studies of child language development*, 1, 75–208.

- Smith, C. S. (1991). *The parameter of Aspect*, vol. 43 of *Studies in Linguistics and Philosophy*. Kluwer, Dordrecht.
- Smith, C. S. (1997). *The parameter of aspect* (Vol. 43). Springer.
- Solomonoff, R. J. (1964). A formal theory of inductive inference. Part I. *Information and control*, 7(1), 1–22.
- Steels, Luc. (2011). *Design Patterns in Fluid Construction Grammar*. John Benjamins Publishing Company.
- Strand, O. (2012, October 13). Tense and Aspect in DRT and SDRT.
- Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, 30(3), 299–312.
- Talmy, L. (2000). *Toward a cognitive semantics* (Vol. 2). MIT press Cambridge.
- Tulving, E. (1983). *Elements of episodic memory*. Clarendon Press Oxford.
- Van Benthem, J. F. A. . (1983). *The Logic of Time. A Model-Theoretic Investigation into the Varieties of Temporal Antology and Temporal Discourse*.
- Vendler, Z. (1967). *Linguistics in philosophy*. Cornell University Press Ithaca.
- Verkuyl, H. J. (1996). *A theory of aspectuality: The interaction between temporal and atemporal structure* (Vol. 64). Cambridge University Press.
- Verkuyl, H. J., & Loux-Schuringa, J. A. (1985). Once upon a tense. *Linguistics and Philosophy*, 8(2), 237–261.
- Vikner, S. (1985). Reichenbach revisited: one, two, or three temporal relations? *Acta Linguistica Hafniensia*, 19(2), 81–98.

14 Annexes

14.1 Listes d'exemples et de leurs interprétations

Voir chapitre [7.4 Lecture des arbres de AIS](#) page 148 pour les explications des interprétations. Certains énoncés considérés comme corrects peuvent être "bizarres", nous précisons alors dans les commentaires le contexte nécessaire. Tout les énoncés présentés ici sont traités par notre modèle.

14.1.1 Exemples avec l'imparfait

	Énoncé	Répétition "action"	Répétition de l'ensemble de la situation	Ensemble de la période	Période déterminée	Attitude	Commentaires
1	<i>Marie mangeait du gâteau en été.</i>		Situation répétée	Ensemble	Déterminée	<i>manger du gâteau</i>	On parle d'été en opposition aux autres saisons. Uniquement possible si le contexte fournit une période de référence assez grande (à cette époque...)
		Action répétée	Situation unique	Ensemble	Déterminée	<i>manger du gâteau</i>	Un seul été (l'an dernier par ex.)
2	<i>Marie mangeait du gâteau en trois mois.</i>		Situation répétée	Ensemble	Non-déterminée	<i>manger du gâteau</i>	Interprétation inchoative (<i>Marie commençait à manger du gâteau au bout d'un certain temps</i>) ! Uniquement possible si le contexte fournit une période de référence assez grande (à cette époque...)
3	<i>Marie mangeait du gâteau en trois heures.</i>		Situation répétée	Ensemble	Non-déterminée	<i>manger du gâteau</i>	Interprétation inchoative (<i>Marie commençait à manger du gâteau au bout d'un certain temps</i>) ! Uniquement possible si le contexte fournit une période de référence assez grande (à cette époque...)
4	<i>Marie mangeait du gâteau en trois minutes.</i>		Situation répétée	Ensemble	Non-déterminée	<i>manger du gâteau</i>	Interprétation inchoative (<i>Marie commençait à manger du gâteau au bout d'un certain temps</i>) ! Uniquement possible si le contexte fournit une période de référence assez grande (à cette époque...)

5	Marie mangeait du gâteau pendant trois mois.						Incorrect
6	Marie mangeait du gâteau pendant trois heures.	Action unique	Situation répétée	Ensemble	Non-déterminée	Ensemble de l'énoncé	Uniquement correcte si le contexte fournit une période supérieure à trois heures (à cette époque)
7	Marie mangeait du gâteau pendant trois minutes.	Action unique	Situation répétée	Ensemble	Non-déterminée	Ensemble de l'énoncé	Uniquement correcte si le contexte fournit une période supérieure (à cette époque)
8	Marie mangeait du gâteau pendant cet été.						Incorrect
9	Marie mangeait du gâteau pendant ces trois mois.	Action répétée		Ensemble	Déterminée	manger du gâteau	Incomplet si contexte absent (on finit sur un ground).
		Action répétée	Situation répétée	Ensemble	Déterminée	manger du gâteau	On parle alors de trois mois spécifiques (tous les juin-juillet-août)
10	Marie mangeait du gâteau pendant ces trois heures.			Ensemble	Déterminée	Ensemble de l'énoncé	Incomplet si contexte absent (il faut préciser une période en apposition, par exemple : hier. Ou continuer sur une narration)
			Situation répétée	Ensemble	Déterminée	Ensemble de l'énoncé	Incomplet si contexte absent (il faut préciser une période en apposition, par exemple : à cette époque. Ou continuer sur une narration) On parle alors de trois heures spécifiques (entre midi et trois heures)
11	Marie mangeait du gâteau pendant ces trois minutes.	Action unique	Situation répétée	Ensemble	Déterminée	Ensemble de l'énoncé	Incomplet si contexte absent (on finit sur un ground). On parle alors de trois minutes spécifiques (pendant le journal météo)
12	Marie courait un marathon en été.	Action unique	Situation répétée	Partielle	Non-déterminée	Ensemble de l'énoncé	
13	Marie courait un marathon en trois mois.			Partielle	Déterminée	courir un marathon	Un est quantitatif. Le contexte doit préciser une période supérieure à trois heures(à cette époque)
14	Marie courait un marathon en trois heures.	Action unique	Situation répétée	Ensemble	Non-déterminée	courir un marathon	Un marathon a le sens de une distance de 42 km. Le un n'est pas quantitatif. Le contexte doit préciser une période supérieure à trois heures(à cette époque)
15	Marie courait un marathon						Incorrect

	<i>en trois minutes.</i>	
16	<i>Marie courait un marathon pendant trois mois.</i>	Incorrect
17	<i>Marie courait un marathon pendant trois heures.</i>	Incorrect
18	<i>Marie courait un marathon pendant trois minutes.</i>	Incorrect
19	<i>Marie courait un marathon pendant cet été.</i>	Incorrect
20	<i>Marie courait un marathon pendant ces trois mois.</i>	Uniquement correct dans le cadre d'une narration, s'il se passe quelque chose pendant l'action (le modèle finit avec un ground), et si ces <i>trois mois</i> sont déterminés et en apposition.
21	<i>Marie courait un marathon pendant ces trois heures.</i>	Uniquement correct dans le cadre d'une narration, s'il se passe quelque chose pendant l'action (le modèle finit avec un ground), et si ces <i>trois heures</i> sont déterminés et en apposition.
22	<i>Marie courait un marathon pendant ces trois minutes.</i>	Uniquement correct dans le cadre d'une narration, s'il se passe quelque chose pendant l'action (le modèle finit avec un ground), et si ces <i>trois minutes</i> sont déterminés et en apposition.

14.1.2 Exemples avec le passé composé

	Énoncé	Répétition "action"	Répétition de l'ensemble de la situation	Ensemble de la période	Période déterminée	Attitude	Commentaires
23	<i>Marie a mangé du gâteau en été.</i>	Action unique		Partielle	Déterminée	<i>manger du gâteau</i>	
		Action répétée		Entière	Déterminée	<i>Manger (plusieurs fois) du gâteau</i>	
		Action répétée		Entière	Non-déterminée	<i>manger du gâteau</i>	lorsque on traite <i>été</i> comme une période générique (undetermined) la répétition de manger du gâteau est obligatoire.
24	<i>Marie a mangé du gâteau en</i>	Action		Partielle	Déterminée	<i>manger du</i>	<i>trois mois</i> déterminés. L'action est exceptionnelle,

	trois mois.	prédiquée				gâteau	interdite, ou une réfutation. → Peu plausible
25	Marie a mangé du gâteau en trois heures.	Action prédiquée		Partielle	Déterminée	manger du gâteau	trois heures déterminées. L'action est exceptionnelle, interdite, ou une réfutation. → Peu plausible
		Action unique		Entière	Non-déterminée	manger du gâteau	Shift : Au bout de trois heures, Marie a commencé à manger du gâteau.
26	Marie a mangé du gâteau en trois minutes.	Action prédiquée		Partielle	Déterminée		Pendant trois minutes déterminées. L'action est exceptionnelle, interdite, ou une réfutation. → Peu plausible
		Action unique		Entière	Non-déterminée		Shift : Au bout de trois minutes, Marie a commencé à manger du gâteau.
27	Marie a mangé du gâteau pendant trois mois.	Action répétée		Entière	Non-déterminée		
28	Marie a mangé du gâteau pendant trois heures.	Action unique		Entière	Non-déterminée		
29	Marie a mangé du gâteau pendant trois minutes.	Action unique		Entière	Non-déterminée		
30	Marie a mangé du gâteau pendant cet été.	Action unique		Partielle	Déterminée		
		Action répétée		Entière	Déterminée		
31	Marie a mangé du gâteau pendant ces trois mois.	Action unique		Partielle	Déterminée		
		Action répétée		Entière	Déterminée		
32	Marie a mangé du gâteau pendant ces trois heures.	Action unique		Entière	Déterminée		
33	Marie a mangé du gâteau pendant ces trois minutes.	Action unique		Entière	Déterminée		
34	Marie a couru un marathon	Action unique		Partielle	Déterminée		

	<i>en été.</i>						
35	<i>Marie a couru un marathon en trois mois.</i>			Partielle	Déterminée		Uniquement possible si les <i>trois mois</i> sont déterminés et <i>un</i> est quantitatif
36	<i>Marie a couru un marathon en trois heures.</i>	Action unique		Entière	Non-déterminée		
37	<i>Marie a couru un marathon en trois minutes.</i>	Incorrect					
38	<i>Marie a couru un marathon pendant trois mois.</i>	Incorrect					
39	<i>Marie a couru un marathon pendant trois heures.</i>	Action unique		Entière	Non-déterminée		Uniquement possible si le marathon n'a pas été achevé (le module perceptif doit accepter que <i>courir un marathon</i> soit atélique)
40	<i>Marie a couru un marathon pendant trois minutes.</i>	Action unique		Entière	Non-déterminée		Uniquement possible si le marathon n'a pas été achevé (le module perceptif doit accepter que <i>courir un marathon</i> soit atélique)
41	<i>Marie a couru un marathon pendant cet été.</i>	Action unique		Partielle	Déterminée		
42	<i>Marie a couru un marathon pendant ces trois mois.</i>	Action unique		Partielle	Déterminée		
43	<i>Marie a couru un marathon pendant ces trois heures.</i>	Action unique		Entière	Déterminée		
44	<i>Marie a couru un marathon pendant ces trois minutes.</i>	Incorrect					

14.1.3 Avec apposition

	Énoncé	Répétition "action"	Répétition de l'ensemble de la situation	Ensemble de la période	Période déterminée	Attitude	Commentaires
45	<i>Cet été, Marie mangeait du</i>	Action répétée		Entière	Déterminée		

	<i>gâteau.</i>						
46	<i>Cet été, Marie courait un marathon.</i>	Incorrect, sauf comme début de narration.					
47	<i>Cet été, Marie a mangé du gâteau.</i>	Action répétée		Entière	Déterminée		
		Action unique		Partielle	Déterminée		
48	<i>Cet été, Marie a couru un marathon.</i>	Action unique		Partielle	Déterminée		
49	<i>Ce matin⁵¹, Marie mangeait du gâteau.</i>	Action répétée		Entière	Déterminée		
50	<i>Ce matin, Marie courait un marathon.</i>	Incorrect, sauf comme début de narration.					
51	<i>Ce matin, Marie a mangé du gâteau.</i>	Action répétée		Entière	Déterminée		
		Action unique		Partielle	Déterminée		
52	<i>Ce matin, Marie a couru un marathon.</i>	Action unique		Partielle	Déterminée		

51 On considérera dans les exemples 49 à 52 que l'ordre de grandeur de *matin* est supérieur à celui de *manger du gâteau* et *courir un marathon*.

14.2 Extraits de l'implémentation

14.2.1 syn_file.pl

Création des règles syntaxiques. 'drule' signifie que la règle est dynamique (par opposition au static de prolog).

```
init_syn :-
    assert(drule(a_ip, [a_tp, a_pp], 1)),
    assert(drule(a_ip, [a_tp], 1)),
    assert(drule(a_tp, [a_vp, ifl], 2)),
    assert(drule(a_vp, [a_vp, a_pp], 1)),
    assert(drule(a_vp, [vrb, a_pp], 1)),
    assert(drule(a_vp, [vrb], 1)),
    assert(drule(a_pp, [cj, a_dp], 1)),
    assert(drule(a_pp, [a_dp], 1)),
    assert(drule(a_dp, [sn], 1)),
    assert(drule(a_dp, [d, np], 1)),
    assert(drule(s, [a_dp, a_ip], 2)),
    true.
```

14.2.2 per_file.pl

Création des données perceptives. Le 'd' dans 'per_dunit' indique une unité dynamique, comme pour les règles syntaxiques.

Le préfixe 'i_' indique l'identifiant d'une image perceptive. Les autres valeurs sont décrites un peu plus bas, dans [14.2.4 tmt_perception.pl](#).

```
init_per :-
    assert(per_dunit(i_year, 0, 6, 3)),
    assert(per_dunit(i_month, 0, 5, 3)),
    assert(per_dunit(i_day, 0, 4, 3)),
    assert(per_dunit(i_hour, 0, 3, 3)),
    assert(per_dunit(i_minute, 0, 2, 3)),
    assert(per_dunit(i_spectacle, 1, 3, 3)),
    assert(per_dunit(i_long_spectacle, 1, 4, 3)),
    assert(per_dunit(i_1914, 1, 6, 3)),
    assert(per_dunit(i_1939, 1, 6, 3)),
    assert(per_dunit(i_2010, 1, 6, 3)),
    assert(per_dunit(i_2013, 1, 6, 3)),
    assert(per_dunit(i_ronfler, 1, 3, 2)),
    assert(per_dunit(i_ecrire_lettre, 1, 3, 3)),
    assert(per_dunit(i_ecrire_article, 1, 5, 3)),
```

```

assert(per_dunit(i_ecrire_livre, 1, 6, 3)),
assert(per_dunit(i_manger_cantine, 1, 3, 3)),
assert(per_dunit(i_manger_pomme, 1, 2, 3)),
assert(per_dunit(i_manger_du_gateau, 1, 3, 2)),
assert(per_dunit(i_manger_le_gateau, 1, 3, 1)),
assert(per_dunit(i_ouvrir_le_coffre, 1, 2, 3)),
assert(per_dunit(i_atteindre_le_sommet, 1, 1, 1)),
assert(per_dunit(i_achever_travail, 1, 1, 3)),
assert(per_dunit(i_interrompre_travail, 1, 1, 3)),
assert(per_dunit(i_boire_du_champagne, 1, 3, 2)),
assert(per_dunit(i_day_after, 1, 4, 3)),
assert(per_dunit(i_day_before, 1, 4, 3)),
assert(per_dunit(simul, 0, 0, 0)),
assert(per_dunit(before, 0, 0, 0)),
true.

```

14.2.3 sem_file.pl

Création des données "sémantiques", manipulées par le module aspectuel. Le 'd' dans 'dsem_unit' indique une unité dynamique, comme pour les règles syntaxiques.

Le premier attribut est un identifiant, le second attribut fournit la liste des valeursinstanciées de l' AIS correspondant.

```

init_sem :-
  assert(dsem_unit(s_fr_passer, [im:i_passer|_])),
  assert(dsem_unit(s_fr_sonner, [im:i_sonner|_])),
  assert(dsem_unit(s_fr_declencher, [im:i_declencher|_])),
  assert(dsem_unit(s_fr_aboyer, [im:i_aboyer|_])),
  assert(dsem_unit(s_fr_chanter, [im:i_chanter|_])),
  assert(dsem_unit(s_fr_tirer, [im:i_tirer|_])),
  assert(dsem_unit(s_fr_apres, [im:a_after, vp:f|_])),
  assert(dsem_unit(s_fr_avant, [im:im:a_before, vp:f|_])),
  assert(dsem_unit(s_xx_later, [im:a_after, vp:f|_])),
  assert(dsem_unit(s_fr_fin, [im:a_after, vp:f|_])),
  assert(dsem_unit(s_xx_spectacle, [im:i_spectacle|_])),
  assert(dsem_unit(s_xx_long_spectacle, [im:i_long_spectacle|_])),
  assert(dsem_unit(s_xx_repas, [im:i_repas|_])),
  assert(dsem_unit(s_null, [_])),
  assert(dsem_unit(s_fr_manger, [im:i_manger|_])),
  assert(dsem_unit(s_fr_derniere, [im:a_last|_])),
  assert(dsem_unit(s_xx_morning, [im:i_morning|_])),
  assert(dsem_unit(s_fr_tousser, [im:i_tousser|_])),
  assert(dsem_unit(s_fr_discuter, [im:i_discuter|_])),
  assert(dsem_unit(s_fr_payer, [im:i_payer|_])),
  assert(dsem_unit(s_fr_etre_tense_futursimple, [im:i_etre|_])),
  assert(dsem_unit(s_fr_en_train_de, [im:a_simul, vp:g|_])),
  assert(dsem_unit(s_fr_entrer, [im:i_entrer|_])),
  assert(dsem_unit(s_fr_etre_tense_passesimple, [im:i_etre, vp:f|_])),
  assert(dsem_unit(s_fr_article_undet, [an:u|_])),
  assert(dsem_unit(s_fr_peindre, [im:i_peindre|_])),
  assert(dsem_unit(s_fr_plusieurs, [cnt:0|_])),
  assert(dsem_unit(s_fr_quelques, [cnt:0|_])),
  assert(dsem_unit(s_fr_ses, [cnt:0, det:l|_])),

```

```

assert(dsem_unit(s_fr_son, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_sa, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_cette, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_cet, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_ce, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_ces, [cnt:0, det:1|_])),
assert(dsem_unit(s_fr_les, [cnt:0, det:1|_])),
assert(dsem_unit(s_fr_le, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_la, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_l, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_un, [cnt:1|_])),
assert(dsem_unit(s_fr_un, [cnt:1|_])),
assert(dsem_unit(s_fr_une, [cnt:1|_])),
assert(dsem_unit(s_fr_deux, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_trois, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_dix, [cnt:1, det:1|_])),
assert(dsem_unit(s_fr_des, [cnt:0, det:0|_])),
assert(dsem_unit(s_fr_du, [cnt:0, det:0|_])),
assert(dsem_unit(s_fr_de_la, [cnt:0, det:0|_])),
assert(dsem_unit(s_fr_voyage, [im:i_voyage|_])),
assert(dsem_unit(s_fr_cj_a, [an:a|_])),
assert(dsem_unit(s_fr_ecrire, [im:i_ecrire|_])),
assert(dsem_unit(s_xx_book, [im:i_book|_])),
assert(dsem_unit(s_xx_books, [im:i_book|_])),
assert(dsem_unit(s_xx_letters, [im:i_letter|_])),
assert(dsem_unit(s_xx_letter, [im:i_letter|_])),
assert(dsem_unit(s_xx_2010, [im:i_2010, an:a|_])),
assert(dsem_unit(s_fr_ecrire_lettre_det, [im:i_ecrire_lettre, an:a|_])),
assert(dsem_unit(s_fr_ecrire_lettre, [im:i_ecrire_lettre|_])),
assert(dsem_unit(s_fr_ecrire_livre, [im:i_ecrire_livre|_])),
assert(dsem_unit(s_fr_ecrire_article, [im:i_ecrire_article|_])),
assert(dsem_unit(s_fr_manger_cantine, [im:i_manger_cantine|_])),
assert(dsem_unit(s_xx_minute, [im:i_minute, an:u|_])),
assert(dsem_unit(s_xx_winter, [im:i_winter|_])),
assert(dsem_unit(s_fr_manger_cette_pomme, [im:i_manger_pomme, an:a|_])),
assert(dsem_unit(s_fr_manger_ces_pommes, [im:i_manger_pomme, an:a|_])),
assert(dsem_unit(s_xx_dhour, [im:i_hour, an:a|_])),
assert(dsem_unit(s_xx_dyear, [im:i_year, an:a|_])),
assert(dsem_unit(s_fr_pendant, [im:simul, vp:g|_])),
assert(dsem_unit(s_fr_dans, [im:after, an:u, pa:np|_])),
assert(dsem_unit(s_fr_depuis, [im:before, pa:p|_])),
assert(dsem_unit(s_xx_year, [im:i_year, an:u|_])),
assert(dsem_unit(s_xx_hour, [im:i_hour, an:u|_])),
assert(dsem_unit(s_fr_quand, [_])),
assert(dsem_unit(s_fr_ronfler, [im:i_ronfler|_])),
assert(dsem_unit(s_fr_gateau, [im:i_gateau|_])),
assert(dsem_unit(s_fr_pomme, [im:i_pomme|_])),
assert(dsem_unit(s_fr_manger_du_gateau, [im:i_manger_du_gateau|_])),
assert(dsem_unit(s_fr_manger_le_gateau, [im:i_manger_le_gateau|_])),
assert(dsem_unit(s_fr_manger_du_gateau2, [im:i_manger_du_gateau, vp:g, an:u|_])),
assert(dsem_unit(s_xx_2013, [im:i_2013, an:a|_])),
assert(dsem_unit(s_fr_ouvrir_le_coffre, [im:i_ouvrir_le_coffre|_])),
assert(dsem_unit(s_xx_now, [an:a|_])),
assert(dsem_unit(s_fr_disparaitre, [im:i_disparaitre|_])),
assert(dsem_unit(s_fr_prendre_bijoux, [im:i_prendre_bijoux|_])),
assert(dsem_unit(s_fr_mourir, [im:i_mourir|_])),
assert(dsem_unit(s_fr_atteindre_le_sommet, [im:i_atteindre_le_sommet|_])),
assert(dsem_unit(s_fr_achever_travail, [im:i_achever_travail|_])),
assert(dsem_unit(s_fr_interrompre_travail, [im:i_interrompre_travail|_])),
assert(dsem_unit(s_fr_boire_du_champagne, [im:i_boire_du_champagne|_])),
assert(dsem_unit(s_fr_en, [vp:f|_])),
assert(dsem_unit(s_fr_tense_present, [di:nr, pa:np|_])),
assert(dsem_unit(s_fr_tense_imparfait, [vp:g, di:fr, pa:p|_])),

```

```

assert(dsem_unit(s_fr_tense_participepasse, [vp:f,di:fr,pa:p|_])),
assert(dsem_unit(s_fr_tense_futursimple, [di:fr,pa:np|_])),
assert(dsem_unit(s_fr_tense_passesimple, [vp:f,di:fr,pa:p|_])),
assert(dsem_unit(s_fr_tense_passecompose, [vp:f,di:nr,pa:p|_])),
assert(dsem_unit(s_fr_demain, [im:i_day_after,pr:s_xx_now,pa:np,an:a|_])),
assert(dsem_unit(s_fr_aujourd'hui, [im:i_day,pr:s_xx_now,pa:np,an:a|_])),
assert(dsem_unit(s_fr_hier, [im:i_day_before,pr:s_xx_now,pa:p,an:a|_])),
assert(dsem_unit(s_fr_lendemain, [im:i_day_after,pa:np,an:a|_])),
assert(dsem_unit(s_xx_day, [im:i_day|_])),
assert(dsem_unit(s_fr_jour, [im:i_day,vp:f|_])),
assert(dsem_unit(s_fr_journee, [im:i_day,vp:g|_])),
assert(dsem_unit(s_xx_month, [im:i_month,an:u|_])),
assert(dsem_unit(s_xx_august, [im:i_august|_])),
true.

```

14.2.4 tmt_perception.pl

Voici le formatage des informations perceptives telle qu'il a été implémenté. L'implémentation est complètement arbitraire et n'est donnée qu'à titre indicatif. "Type" correspond à la valeur télicité.

```

%
%-----
%% Perception Unit
%%
%% Elem      - Élément (Approche ou image perceptuelle)
%% Image     - L'élément est-il une image perceptuelle (0/1) ?
%% Magnitude - Ordre de grandeur du phénomène
%% Type      - Les différentes configurations d'homogénéité
%%           0 - approche
%%           1 - insécable
%%           2 - auto-similaire
%%           3 - (1+2)
%%           9 - non-temporel ("pomme")
%%
%
%-----
%
%-----
%% per_unit(?Elem, ?Image, ?Magnitude, ?Type)
%
%-----

per_unit(Elem, Image, Magnitude, Type) :-
    per_dunit(Elem, Image, Magnitude, Type).
%% Approches

```

```

%% Elles se distinguent des images par un ordre de grandeur nul.
%% On se sert ici des foncteurs statiques de tmt_ss.pl.
per_unit(A, 0, 0, 0) :-
    approach(A),!.
%% per_tunit correspond à une unité temporaire de la perception,
%% qui ne sera pas stockée dans les bases.
per_unit(Elem, Image, Magnitude, Type) :-
    per_tunit(Elem, Image, Magnitude, Type).

[...]

```

14.2.5 tmt_model.pl

Algorithme de la fusion aspectuelle en prolog. De nombreux attributs sont uniquement utilisés à des fins de traçage ("debug").

```

[...]

%
%#####
%%
%% MAIN SEMANTIC MERGE
%%
%% Chaque semt_merge est une fusion différente et donc exclusive.
%% Par convention le SS B est le dominant (syntaxiquement
parlant)
%%
%#####

%%
-----
%% semt_merge(+AIS_Comp_before, +AIS_Head_before,
%%           -AIS_Comp_after, -AIS_Head_after,
-AIS_Result_after,
%%           -LogList, -Success, -AIS_predicated,
%%           -AIS_infosA, -AIS_InfosB)
%%
%% Merge de deux AIS (AIS_Comp_before et AIS_Head_before).
%% AIS_Comp_after et AIS_Head_after correspondront à leur état
%% après le merge, tandis que AIS_Result_after est le AIS
%% résultant du merge.
%%
%% LogList stock les infos de logs du merge (voir tmt.spy.pl).
%% Success indique le succès (1) ou non (<0) du merge.
%% AIS_predicated indique le AIS prédiqué (s'il y en a un).

```

```

%% AIS_infos donne des infos sur les AIS merges (on les retrouve
%% dans LogList).
%%
%% nt : non temporel
%% tmp : temporel
%%
-----

sem_t_merge(A, B, A1, B1, R, L, S, P, IA, IB) :-
  dspv(L, "<p class='title'>predication</p>", S),
  sem_t_predication(A, A0, L, S, P, IA),
  sem_t_predication(B, B0, L, S, P, IB),
  log_predication(L, P),
  dspv(L, "<p class='title'>ajustements</p>", S),
  ajustements(A0, B0, A1, B1, L, P, S, IA, IB),
  %% Merges non-temporels. À faire avant le merge perceptif
  ---
  dspv(L, "<p class='title'>determination merge</p>", S),
  sem_tntp_merge(det, A1, B1, R, L, S),
  dspv(L, "<p class='title'>countable merge</p>", S),
  sem_tntp_merge(cnt, A1, B1, R, L, S)
  %%
  -----
  dspv(L, "<p class='title'>image merge</p>", S),
  sem_t_image_merge(A1, B1, R, L, S),          %% nt
  %%
  -----
  dspv(L, "<p class='title'>anchoring merge</p>", S),
  sem_t_anchoring_merge(A1, B1, R, L, S),
  dspv(L, "<p class='title'>viewpoint merge</p>", S),
  sem_t_viewpoint_merge(A1, B1, R, L, S),
  dspv(L, "<p class='title'>distance merge</p>", S),
  sem_t_distance_merge(A1, B1, R, L, S),
  dspv(L, "<p class='title'>past/non-past merge</p>", S),
  sem_t_past_merge(A1, B1, R, L, S),
  %%
  sem_t_merge_success(S).

%
%-----
%% sem_t_predication/4
%
%-----

%% Image sans prédication
sem_t_predication(SS, SS, _, _, _, _).
%% Prédication

```

```

sem_t_predication(SSi, SSo, _, _, P, I) :-
    %% On vérifie qu'aucune prédication n'a été faite
    var(P),
    %% Image requise (donc ni approche, ni durée).
    im(SSi, Im), per_unit2(Im, 1, _, _),
    %%
    vp(SSi, g), an(SSi, _),
    vp(SSo, f), an(SSo, a),
    %% alternative, sans contrainte sur le viewpoint
    %an(SSi, _),
    %an(SSo, a),
    %% alternative, avec contraintes sur vp ET determination
    %vp(SSi, g), an(SSi, a),
    %vp(SSo, f), an(SSo, a),
    %%
    lcopy_list(SSi, SSo),
    P = Im,
    %% Pour les infos de log
    checkE(predication:Im, I).

```

[...]