Chapitre 4

Un dispositif ou le dur désir de représenter

C'est à l'étude des problèmes de représentation et construction du sens des configurations du type `moulin à N` dans une analyse automatique qu'est consacré la première partie de notre travail. Celui-ci vise donc à construire un dispositif pour représenter matériellement les phénomènes complexes en jeu pour une modélisation des connaissances de telles expressions et en particulier les phénomènes décrivant la construction du sens de ces expressions. Notre approche se situe dans le cadre d'une représentation des connaissances et d'une construction du sens des fragments de langue étudiés par un affinage progressif de ces mécanismes de représentation et de construction.

Dans les approches retenues pour la construction d'un dispositif pour l'étude des séquences `moulin à N2`, notre propos sera le suivant :

> Un dispositif pour une analyse automatique des formes `moulin à N2` doit identifier, classer et décrire ces formes pour construire une taxonomie ouverte et évolutive de cette classe particulière de composés.

On doit donc étudier les procédures de traitements des rapports forme-sens qui sont à l'oeuvre dans les dispositifs informatiques construits : comment à partir de représentations statiques du lexique, mettre en place des opérations dynamiques pour l'interprétation?

4.1. Préambule : Le dispositif est au bord d'un gouffre

Du désir de représenter l'excès dans la langue au possible de représentation dans un dispositif

Sommes-nous à même aujourd'hui de représenter le langage naturel? La réponse est clairement non. Des représentations locales, partielles sont possibles, elles ne sont pas à même de prendre la mesure de l'étendue des possibles dans les énoncés. Si les faits de langue maximalisent les informations, la finesse de ces informations, leurs connexions, ..., leurs représentations tendent malgré tout à les réduire.

La langue use ses représentations

La langue évolue : construire des représentations matérielles du langage naturel, c'est donc être contraint à réviser en permanence les représentations. De plus, à un moment donné on n'en sait jamais assez sur un fait de langue; notre savoir évolutif sur la langue peut donc conduire à modifier les résultats initiaux. De même, le développement de techniques nouvelles peut aussi amener à interroger les outils utilisés pour la représentation.

Le complexe de la langue n'est pas convertible dans "une monnaie" de représentation à "taux fixe"

Il faut donc s'attacher à des esquisses de représentation. Qu'il s'agit ensuite de faire évoluer : "dégrossir les esquisses". Ce qui est représenté aujourd'hui, peut être à re-représenter demain. De même pour la représentation de la construction du sens : ce qui sensé aujourd'hui peut devenir non-sens demain.

Le fossé entre le complexe de la langue et les outils pour sa représentation est considérable

La principale difficulté à laquelle est confronté le travail de représentation des faits de langue tient dans le rapport inégal qui existe entre d'une part, la complexité des savoirs qui agissent dans ces faits de langue et d'autre part, les outils disponibles pour la mise en place d'une telle représentation. Les énoncés linguistiques réalisent une imbrication des différents types de savoirs qu'ils véhiculent. Leur articulation est partie intégrante de la structure de ces énoncés. La richesse, la finesse des savoirs qui participent dans un fait de langue ne se laissent pas appréhender facilement. On est donc loin d'être en mesure de représenter ce qu'un être humain est capable de faire dans son activité de production d'énoncés. Vouloir représenter une telle complexité de phénomènes met en lumière la nécessité de disposer d'outils adéquats pour saisir les interactions de savoirs qui se mettent en place dans les énoncés visés. Les outils informatiques existants ne permettent pas encore de prendre la mesure de la réalité complexe d'un fait de langue. De plus, comme ces outils pour la représentation matérielle du langage naturel ne nous sont pas fournis de "toutes pièces", il reste à les construire, tout au moins à essayer de le faire. Et c'est dans la réalisation de cette construction d'outils pour la représentation que l'on prend la mesure du gouffre qui existe entre l'objet à représenter et les moyens à mettre en oeuvre pour cette représentation.

Le dispositif ou l'impatience de représentation

Un dispositif pour le TALN ne peut donc pas encore réaliser la représentation de l'excès de la langue, il en reste au possible du représentable.

4.2. Un dispositif qui construit, utilise et modifie des hiérarchies

Un dispositif qui privilégie un cadre de représentation des connaissances sous forme d'objets dans une taxonomie de classes a du mal à prendre en compte le fait que dans un domaine particulier où les connaissances ne sont pas encore clairement établis, il faut faire évoluer les représentations initiales pour construire pas à pas un modèle possible. Si les éléments constitutifs d'un modèle de représentation sont organisés de manière hiérarchique, il est souvent difficile de modifier cette hiérarchie quand les observations révélées par l'expérience révèlent une telle évolution[1]. La modélisation d'un domaine de connaissance sous forme d'une hiérarchie de classe se révèle délicate pour représenter un domaine où les connaissances sont peu (ou ne sont pas) organisées de manière définitive[2].

4.2.1. De la complexité de l'objet à représenter

S'il s'agit de représenter les éléments constitutifs d'une analyse automatique pour la construction du sens, on est confronté à une masse de phénomènes. Les interférences lexicales, syntaxiques, sémantiques, contextuelles sont nombreuses et complexes entre les différents mécanismes qui réalisent la construction du sens d'un énoncé.

4.2.2. Organiser le savoir lexical

Le début des années 80 a vu la montée de formalismes qui ont donné une place centrale au lexique, et par conséquent à des structures de traits complexes, et à l'unification sur ces structures. L'utilisation de tels formalismes pose clairement le problème de représenter les articulations forme-sens qui se jouent au niveau des éléments lexicaux : penser l'organisation lexicale. Une organisation hiérarchique avec héritage de propriétés pose de nombreux problèmes en raison justement de la complexité des phénomènes à représenter. Dans notre domaine particulier des séquences `moulin à N2`, la construction d'une telle hiérarchie n'est pas triviale et pousse à réfléchir sur l'organisation des connaissances dans un tel cadre. Là encore il s'agit de bien comprendre que ce n'est pas tant de construire une hiérarchie qui importe le plus (ce n'est d'ailleurs pas le plus difficile, avec la programmation objet tout est objet, i.e. on peut ainsi tout représenter dans une taxonomie de classes hiérarchisées), mais bien de donner les moyens de faire évoluer cette hiérarchie. Plus précisément, dans notre domaine restreint, l'interprétation du patron `moulin à N2` met en lumière les difficultés à organiser le savoir sur des formes qui autorisent des extensions de sens : faut-il pour cela considérer que ces extensions de sens doivent être prises en compte dans une hiérarchie de concepts, ou bien dans une hiérarchie lexicale? De plus quel est le lien de dépendance entre ces types de hiérarchies? Les entrées lexicales se réduisent-elles à l'addition d'informations héritées d'une hiérarchie conceptuelle?

4.2.3. Du possible dans la représentation

Les problèmes en jeu dans une telle approche mettent donc en avant les difficultés d'organiser le savoir sous forme hiérarchique : si l'on construit une hiérarchie de savoir statique, on a peu de chances de rendre compte des phénomènes en jeu quand les connaissances de ce domaine de

[1] Dans ce cadre, on ne dispose en effet que d'un seul mode d'évolution: la spécialisation (c'est-à-dire la création de sous-classe).

[2] Même si les résultats de la science linguistique sont à jamais provisoires, on peut considérer qu'à un moment donné, certains phénomènes de la langue se prêtent à un classement hiérarchique.

savoir évoluent; si l'on veut construire une hiérarchie qui structure le savoir de manière à préserver l'indétermination ou l'affinage potentiel de certaines représentations, on se heurte à des problèmes complexes. C'est pourtant dans cette deuxième approche de travail qu'il faut se situer si l'on veut espérer modéliser les domaines de savoir évolutif du langage puis construire des dispositifs cohérents capables de gérer ces évolutions. Vouloir représenter, c'est donc réfléchir sur la mise en oeuvre d'outils capables de réaliser les représentations adéquates du domaine visé. Si ces outils tendent à organiser hiérarchiquement les savoirs à représenter, il s'agit tout d'abord de représenter en articulant au mieux les savoirs représentés, puis d'évaluer ce qui dans ce type de représentation, permet de rendre compte des phénomènes visés et enfin de prendre la mesure de ce qu'il reste à faire pour s'en approcher au plus près.

4.2.4. Attacher des scénarios aux entrées lexicales : prédire les possibles d'interprétation

Le travail de Pustejovsky est une illustration de cette volonté de formaliser les éléments lexicaux en mettant en oeuvre des structures qui permettent une compositionalité sémantique. Il s'agit via la notion de *qualia* (inspirée d'Aristote) de regrouper dans cette structure les attributs essentiels d'une entrée lexicale. Même si une entrée lexicale peut être associée à plusieurs interprétations potentielles dans la mesure,*"il est arbitraire de créer des mots de sens différents pour une même entrée lexicale juste parce que ces mots entrent dans des réalisations syntaxiques différentes"* (Pustejovsky 1991). Ce qui est jeu, c'est la possibilité de regrouper au sein d'une structure globale les informations qui permettent de spécifier les différents niveaux de représentation capables de générer et composer le sens. Ce choix de représentation met en place une articulation hiérarchique qui assigne une structure de qualia à chaque entrée lexicale, et qui spécifie les liens entre ces entrées lexicales en pouvant restreindre ces liens sur certains aspects de leur structure lexicale respective (Pustejovsky 1991). Dans le domaine des composés, l'utilisation de la prédicativité nominale peut permettre de restreindre les interprétations possibles des constituants du composé (Habert & Fabre 1993, Fabre 1996). On associe pour cela aux noms à structure argumentale, les mots qui apparaissent dans des scénarios typiques. Il faut malgré tout rester vigilant sur l'identification de tels noms prédicatifs au risque de faire la part trop belle au référentiel et donc d'oublier le contextuel.

4.2.5. Difficulté de l'affinement d'une hiérarchie lexicale en fonction des observations faites lors d'une analyse

4.2.5.1. Les faits de langue sont difficilement réductibles en termes d'objets figés

Représenter dans une taxonomie de classes, c'est se heurter de front au modèle de représentation sous forme d'objet dans lequel "un objet est (principalement) ce qu'il est". Les outils pour la représentation dans une telle approche ont tendance à figer les choses représentées. Et donc quand il s'agit de représenter une matière vivante, les problèmes surgissent rapidement. Si on accepte de penser en termes d'objet, c'est-à-dire admettre que tout est réductible en un ensemble de propriétés définitoires stables, une telle approche se justifie. Malheureusement, on n'en saura jamais assez sur les faits de langue. Le langage crée en permanence. Un fait de langue peut prendre le contre-pied de résultats établis, bousculer les règles définitoires. On est donc contraint à des esquisses de représentation. Reste à savoir si les différentes esquisses en disent toujours plus (et mieux) sur ce qui est représenté.

4.2.5.2. "Donner du mouvement" au jeu de la représentation

S'il apparaît très clairement qu'un choix de représentation qui manipule la notion d'objet dans une taxonomie de classes est adéquate pour fixer un domaine de savoir établi dans une structure hiérarchisée, il s'avère parfois délicat de construire puis de prendre en compte les évolutions de telles représentations hiérarchiques quand celles-ci décrivent des domaines de savoir non stabilisés. Dans une telle approche, la mise en contexte des instances de classe peut révéler la nécessité d'une mise à jour de la structure initialement adoptée pour la définition des classes : celles-ci constituent les moules définitoires dont les caractéristiques seront reprises par les objets dans la relation "une-sorte-de". Dans le cas de la construction d'une hiérarchie lexicale dans laquelle on souhaite inscrire dans la définition des entrées lexicales une articulation forme-sens, une telle représentation doit affronter le problème complexe de la confrontation d'un savoir établi et d'un savoir à venir (non pris en compte par le précédent). La prise en compte de nouvelles informations, de résultats issus de l'analyse, de configurations énonciatives particulières, etc..., peut conduire à modifier, à affiner la représentation initiale. Et c'est justement le travail de mise à jour d'une hiérarchie qui pose problème. Plus on structure un domaine de savoir, plus les liens de dépendance entre les noeuds du graphe d'héritage se complexifient, et il devient ainsi délicat de modifier cette structure hiérarchique. Dans le cadre de notre travail sur la modélisation des configurations nominales du type `moulin à N2`, une telle évolution sur le savoir initial est à l'oeuvre sur, par exemple, des formes du type `moulin à paroles` pour laquelle `moulin` devient `+humain`. Il ne s'agit pas ici prendre le parti de dire que la représentation sous forme hiérarchique est inconcevable; au contraire il s'agit d'adapter les structures hiérarchiques aux processus qui sont à l'oeuvre dans la construction du sens des faits linguistiques.

4.2.5.3. Articuler les représentations

La construction d'une taxonomie ouverte et économique (Milner 1989) doit donc prendre en compte la nécessité de mettre en place des représentations des savoirs qui d'une part articulent ces représentations et d'autre part ne forcent pas les représentations en fonction de rencontres de surface : même si l'on peut rencontrer `"cette 2CV conduit dangereusement"` (G.Gross), il n'est pas souhaitable de classer `"2CV"` dans la classe des `Nhumains`. A l'inverse, c'est justement une représentation qui manipule des mécanismes inférentiels capables de construire, dans ce contexte particulier et temporairement, cette association de `"2CV"` à la classe des `Nhumain` qu'il faut privilégier.

4.2.5.4. Permettre un affinement dans les représentations

Les phénomènes qui participent à la construction du sens d'un énoncé restent encore mal connus. Leur formalisation et leur représentation doivent préserver la non-finitude dans ce qui est formalisé, représenté. Il faut laisser ouvert le champ de travail sur cette modélisation des connaissances qui interagissent dans la construction du sens. En particulier, puisqu'un tel dispositif doit traiter du rapport forme-sens qui est à l'oeuvre dans la construction du sens, il est nécessaire de permettre des articulations entre les savoirs représentés. Comme il s'agit encore de représenter en proposant des esquisses de modélisations pour ces interactions, il est nécessaire que les moyens mis en oeuvre permettent de construire une représentation matérielle de ces articulations de savoirs en ne figeant ni les processus d'interprétation ni les savoirs représentés.

4.2.6. Inscrire des mécanismes inférentiels au sein des représentations

Pour D. Kayser, qui parle de profondeur variable pour la représentation des connaissances, ou pour Pustejovsky qui construit un lexique génératif, il s'agit, dans les deux cas, de construire une hiérarchie de concepts, puis d'attacher aux noeuds de celle-ci des rôles qui guident les inférences éventuelles. On peut d'ailleurs considérer que la construction du sens des expressions `moulin à N2` relèvent de mécanismes d'extension de sens (à la Kayser) ou encore des coercions de type (à la Pustejovsky). Et ces deux cadres révèlent l'inadéquation d'un filtrage statique pour l'étude des séquences `moulin à N`, en effet l'élément en position d'argument ne présente pas les traits nécessaires, ceux-ci doivent lui être ajoutés et au besoin en supprimant des traits initialement présents mais contradictoires (`moulin` devient `+humain` dans `"moulin à paroles"`). Ces deux approches restent cependant éloignées dans les hypothèses initiales suivies. La sémantique lexicale suivie par D. Kayser est d'abord inférentielle. Elle permet la mise de oeuvre de processus interprétatifs capables d'ajuster ou de remodeler les unités manipulées. A l'inverse, l'approche suivie par Pustejovsky impose un point de vue initial particulier sur les unités lexicales représentées sous la forme de structure qualia. Même si certains mécanismes sont définis pour étendre les représentations construites, cette approche s'accommode d'un graphe de types figés qu'il est difficile de faire évoluer.

4.3. Un dispositif qui manipule des prototypes

Notre tâche vise à construire un dispositif dans un cadre de représentation des connaissances qui manipulent des prototypes. La construction de ce dispositif prend appui sur les objectifs suivants :

Ne pas figer les représentations

Dans une approche `Objet(Instance de Classe)` (<- présentation de la PàO), on est contraint de prédéterminer les propriétés pertinentes pour la création des objets : si l'on en sait suffisamment sur un domaine particulier de savoir, une telle approche se révèle généralement adéquate pour représenter ce domaine. Par contre si l'on ne sait pas vraiment quelles sont les propriétés pertinentes des objets que l'on va devoir représenter, une telle approche devient problématique.

S'appuyer sur un savoir minimum établi et affiner la représentation

L'approche `Objet(Prototype)` (<- présentation de la PàP) privilégie une représentation des objets sans référence à un moule initialement prédéterminé. Les objets sont construits "de toutes pièces", en fonction des besoins révélés par le contexte dans lequel cet objet est créé. Un autre `Objet(Prototype)` peut être construit à partir de l'objet initial, il possèdera tout d'abord les mêmes propriétés que celui-ci et pourra ensuite se différencier par ses propres propriétés spécifiques. La notion de prototypes privilégie donc une définition contextuelle des éléments représentés, tout en permettant une évolution de cette représentation suivant les évènements qui obligent à la préciser.

Construire le sens en contexte

Les phénomènes à l'oeuvre dans la construction du sens d'un énoncé révèlent généralement l'importance fondamentale des indices contextuels dans la résolution de l'interprétation de cet énoncé.

- La construction du sens confronte les indices sémantiques dont l'énoncé est porteur : l'interprétation est en effet le résultat de la mise en concurrence des divers indices que le dire parsème dans son trajet de production. Si l'on considère la configuration nominale `moulin à bras` dans un énoncé dont le thème général est centré sur la notion d'énergie, son interprétation sera facilitée par une association de `bras` à `source d'énergie`.

Ces constats conduisent donc à envisager la représentation matérielle de la construction du sens des configurations nominales visées dans un environnement qui, a priori, se prête à une définition contextuelle et évolutive des éléments manipulés. Il s'agit donc dans un premier temps d'identifier le savoir initial minimum qu'il est possible d'attacher aux éléments définis, puis de gérer les évolutions dans la représentation de ce savoir en fonction des résultats que la mise en contexte des éléments définis laisse apparaître. A l'inverse d'une prédiction sur les scénarios possibles que l'on peut attacher aux entrées lexicales et qui limite les interprétations sur le monde fermé du possible, nous souhaitons associer aux entrées lexicales des comportements qui se précisent au regard des évènements contextuels rencontrés.

4.3.1. Construire une analyse automatique guidée par des prototypes

Dans un premier temps, notre tâche est de construire un analyseur syntaxique guidée par une représentation prototypique des différents éléments manipulés (-> partie 2 chapitre 5). Il s'agit ensuite de réaliser un couplage entre analyse syntaxique et analyse sémantique dans un cadre précis de représentation de la construction du sens (celui des configurations nominales de la forme `moulin à X`) (-> partie 2 chapitres 5, 6).

4.3.2. Représentation des connaissances et Objet

4.3.2.1. Des prototypes pour coder la construction du sens de certains composés nominaux

Ce dispositif à approche prototypique illustre les problèmes de représentations des connaissances sous forme prototypique (les notions de catégorie et de prototype seront analysées), puis propose des approches de solution pour la construction du sens des formes `moulin à N2` dans une approche prototypique. Puisque les activités langagières font la part belle à la production d'erreurs (contrôlées), de flou, d'ambiguïté..., et que ces phénomènes sont généralement inscrits dans un contexte particulier, il apparaît cohérent de vouloir s'intéresser à un cadre de représentation des connaissances qui, a priori, offre une grande adaptabilité pour prendre en compte de tels phénomènes. Il s'agit de vérifier si les résultats produits (in)valident cette intuition initiale. Pour illustrer les approches radicalement différentes des dispositifs étudiés, on peut tout de suite noter le grand écart existant entre une approche hiérarchisée et une approche prototypique :

- **Une approche descendante.** La première approche privilégie une représentation sémantique très abstraite pour `moulin`, qu'il faudrait préciser en fonction des besoins rencontrés par le biais de scénarios, de processus inscrits dans la hiérarchie définie.

- **Une approche montante.** Au contraire, la seconde approche met en avant l'utilisation d'"individus" concrets. Il s'agit d'évaluer comment construire un modèle évolutif des connaissances à partir d'un ensemble d'individus concrètement définis et à priori non inscrits dans une hiérarchie prédéterminée.

4.3.2.2. Quels objets linguistiques devons-nous admettre dans notre propre manière de rendre compte des faits ?

> Puisqu'il est difficile d'appréhender globalement les faits de langue examinons au plus près leurs composants.

- Tel mot, tels comportements?

On se trouve donc confronté au problème de définir les individus à représenter. Comment réaliser cette définition prototypique des éléments qui participent à la réalisation des faits de langue. En particulier, quel type de savoir doit-on attacher aux entrées lexicales manipulées : tel mot, tel savoir?

- Au commencement était le mot...

Puisque le sens et la signification sont traditionnellement définis au niveau du mot, nous avons choisi de privilégier une approche de représentation qui associe à chaque entrée lexicale un prototype. D'autant plus qu'il ne s'agit pas ici de dépasser, dans un premier temps, le cadre de la composition nominale, nous en restons en effet à une étude des rapports forme-sens sur "une distance généralement courte". Il convient, malgré tout, de ne pas oublier que certaines

caractéristiques textuelles peuvent influencer le sens particulier des mots (nous y reviendrons infra).

- Percevoir les régularités sur les mots

Notre travail vise donc à associer aux prototypes lexicaux une structure définitoire initiale (qui s'appuiera sur la tradition grammaticale établie).

- Si nous pouvons nous "jouer des mots", que leur représentation essaie d'en tenir compte

Il convient ensuite d'affiner les représentations en fonction du but poursuivi (la reconnaissance et la construction du sens de formes nominales) et des résultats que la mise en contexte de ces prototypes lexicaux révèlent.

4.3.3. Etablir une évaluation critique des pouvoirs expressifs des modèles de représentation

Ce travail vise aussi à établir une comparaison critique du pouvoir expressif de deux cadres de représentation des connaissances : le cadre `Objet(Instance de Classe)` et le cadre `Objet(Prototype)`. Si dans un cas il est a priori plus facile de représenter les savoirs établis, et si dans l'autre une représentation contrainte par les informations contextuelles est a priori plus adéquate, il s'agira d'évaluer les apports pertinents que permet chacun de ces modèles pour la représentation du langage naturel, quand on sait que ce dernier articule très bien ces deux aspects de la manipulation des connaissances.

4.3.4. Représenter les mécanismes de construction du sens

Le dispositif prototypique doit permettre la mise en place d'outils pour la classification des formes `moulin à N2 :` dans ce domaine de savoir où il n'existe pas encore de structuration établie, il s'agit de catégoriser, c'est à dire de construire une hiérarchie en classifiant les données une à une.

4.3.4.1. Représenter en tenant compte des absences du représentable

En particulier, notre dispositif doit permettre de corriger les représentations initiales arbitrairement choisies pour construire des résultats cohérents avec les savoirs déjà validés par notre modèle de construction du sens des formes `moulin à N2`. Ainsi, le dispositif propose dans un premier temps de classer les formes `moulin à N2` de telle sorte que si le résultat produit ne correspond pas à une des facettes sémantiques du modèle initial choisi pour le classement des `moulin à N2`, on puisse intervenir directement sur les représentations initiales des éléments lexicaux pour éventuellement les "faire entrer" dans la taxonomie en cours de construction.

4.3.4.2. Repérer les mécanismes d'extension de sens et (re)construire ou affiner les représentations

Il s'agit dans un deuxième temps de construire des règles pour prendre en compte ces "erreurs" de représentation et pour les ajuster de telle sorte qu'il soit possible de les identifier. Le dispositif doit pouvoir ajuster en supprimant/ajustant des informations si besoin est : ce sera le

cas pour `moulin à sottises`, structure dans laquelle `moulin` devient `+humain`. Nous verrons infra (-> partie 2 chapitre 6) qu'il met en place un filtrage évolutif basé sur la structure d'évènements associés aux noms têtes. Des outils capables de gérer les extensions éventuelles en fonction des rôles hérités sont à définir.

4.3.5. Privilégier une représentation qui permette la spécialisation

4.3.5.1. Construction des composés : une hiérarchie implicite

Puisque une approche prototypique privilégie l'utilisation d'éléments définis contextuellement et modifiables suivant les informations nouvelles rencontrées, nous essaierons de mettre en place un système de traitement qui permette la spécialisation des éléments représentés et construits pour conduire l'analyse. Ainsi, s'il s'agit de construire une configuration nominale du type `N1 à N2` où `N1 = moulin`, l'analyse se déroule en essayant de construire un objet représentant le composé qui ira du plus général (un objet `N1 à N2` sans sémantique définie) au plus particulier (un objet `moulin à N2` avec calcul sémantique intégré). Cette construction par spécialisation est mise en place dans et par les éléments représentés. `Moulin` permet de construire un composé du type `N1 à N2`, et le composé `N1 à N2` se comporte de manière particulière quand le `N1` vaut `moulin`. Dans ce cas on essaie de construire une interprétation de ce composé. Si ce calcul interprétatif réussit, c'est l'objet le plus spécifique qui est considéré comme construit, c'est-à-dire l'objet qui possède une plus grande cohérence de représentation par rapport à ce qu'il représente. Dans le cas contraire, on ne bloque pas la reconnaissance, quitte à produire un résultat non défini ou incomplet (sémantiquement dans notre cas).

4.3.5.2. Des représentations au "jeu" sur les représentations

Il s'agit donc de préserver, au niveau de la représentation puis au niveau de l'implémentation, la possibilité de prendre en compte une évolution à venir dans la mise en place du dispositif. Même si notre travail se limite dans un premier temps à la représentation des problèmes d'interprétation des configurations du type `moulin à N2`, il doit pouvoir, d'une part, ne pas bloquer la reconnaissance de configuration nominale proche de celle-ci, et d'autre part, permettre une prise en compte, à moindre coût, d'éventuelles évolutions de ce modèle initial pour y intégrer la prise en compte de nouvelles représentations supplémentaires (la construction du sens de configurations nominales proches).

4.3.6. Développer des outils de contrôle pour l'analyse

Puisque notre approche se situe dans un cadre de modélisation des connaissances par la construction de représentations qu'il s'agit de faire évoluer, d'affiner, on conçoit aisément qu'il soit nécessaire de disposer d'éléments "dégagés" de ces représentations et qui permettent d'évaluer ce qui est en jeu pour ensuite permettre une prise en compte statique/dynamique des éléments pertinents susceptibles de faire évoluer les représentations initiales. Notre difficulté à isoler les interactions complexes des savoirs à l'oeuvre dans la production d'énoncé accentue la représentation et la construction de tels mécanismes. Il semble malgré tout possible d'y voir plus clair sur une utilisation de processus de contrôle à un méta-niveau si l'on restreint leur étude à des niveaux de complexité moindre. Notre réflexion porte donc aussi sur cet aspect fondamental dans la représentation des connaissances pour un dispositif informatique. D'autant plus que les langages de haut niveau que nous utilisons mettent en place des mécanismes qui favorisent les approches réflexives. Celles-ci peuvent éclairer une étude sur la construction de mécanismes de contrôle pour l'analyse et sur la définition même de ces mécanismes : peut-on construire de tels mécanismes capables de restructurer les représentations

initiales, de modifier les opérations pré-définies... C'est en tout cas ce que nous essaierons de présenter dans notre cadre linguistique choisi.

4.4. Prototypes, représentation des connaissances et modèles du sens

4.4.1. Noyau de sens et prototypes

La notion de prototype semble adéquate pour représenter un noyau de sens tel que celui que C. Fuchs décrit. Il reste à penser les procédures de contrôle pour déterminer et définir les manières dont les chaînes de sens peuvent s'étendre et les directions dans lesquelles elles opèrent. Il convient pour cela de définir des outils qui permettent à partir de représentations de base volontairement sous-déterminée de préciser les informations ainsi représentées. On peut ainsi envisager de systématiser les phénomènes de reprise sur indétermination qui peuvent se produire dans la construction du sens mais que l'on parvient toujours à atténuer ou à éliminer en faisant appel à des connaissances supplémentaires.

En termes de représentation prototypique à la Self (Self Group 1987-1995) par exemple, la définition d'un noyau de sens peut se matérialiser par l'association d'une entrée lexicale et d'une entité prototypique portant ce noyau de sens. Cette dernière structure peut être considérée comme un comportement commun à un ensemble d'unités lexicales qu'il reste à préciser dans certaines configurations interprétatives. Il s'agit donc d'ordonner les comportements sémantiques que l'on peut attacher à une entrée lexicale, c'est-à-dire déterminer les chemins de sens (en termes de comportements particuliers) et leurs directions. Si la notion de prototypes rend possible dans la représentation l'ordonnancement de comportements, il reste à vérifier si les comportements sémantiques d'une entrée lexicale se prêtent à un tel ordonnancement sémantique.

4.4.2. Adaptation d'une représentation prototypique et d'une approche à la Pustejovsky ?

La notion de prototype reste cohérente avec la volonté de capter les régularités lexicales dans les représentations lexicales associées à des mécanismes d'inférence sur ces représentations.

4.4.2.1. Inscrire des processus d'interprétation dans la représentation prototypique

En tant qu'objet, le prototype peut être défini comme un élément qui structure les différents niveaux de représentation des éléments lexicaux. On peut toujours ajouter de l'information dans une représentation prototypique et à moindre coût. Il est toujours possible de faire évoluer les éléments représentés en les spécialisant ou en modifiant leur lien d'héritage : les prototypes n'héritent pas d'un moule structurel définitoire, mais de comportements partagés par un ensemble de prototypes.

L'évolution dans un système de représentation à base de prototypes n'est pas uniquement descendante :

- On peut ajouter et/ou supprimer de l'information dans les prototypes construits.

- On peut inscrire de la délégation[3] de comportements communs et jouer sur ce lien d'héritage : d'une part il est possible de modifier ce lien d'héritage; d'autre part, on peut affiner les comportements partagés et les spécialiser.

Pour une entrée lexicale donnée, il est possible d'associer à la structure définitoire représentant cet élément lexical une suite de comportements différents qu'il est possible d'ordonner (en tenant compte par exemple de leur fréquence de réalisation); il est aussi possible de factoriser ces comportements sur un ensemble d'unités lexicales (peu importe a priori leur "qualité" linguistique, il suffit que l'unité représentée se comporte suivant les processus auxquels elle est attachée).

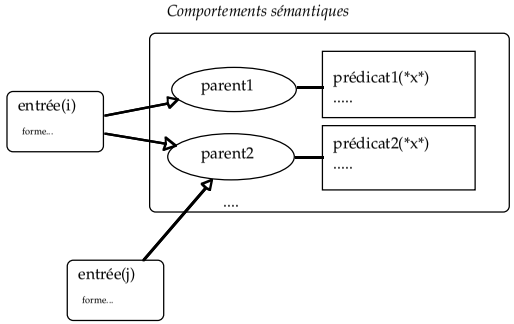


Figure 4.1 : Telle entrée lexicale, tels comportements ?

On peut aussi représenter les unités lexicales en proposant des directions de sens qu'il est possible de suivre en tenant compte des informations rencontrées. S'il y a blocage dans l'application de tel comportement sémantique donné, il faut pouvoir récupérer un chemin de sens adéquat pour la configuration visée. Dans l'exemple de `"fast"` traité par Pustejovsky, on peut inscrire dans le comportement sémantique de l'adjectif qu'il peut modifier le prédicat attaché au rôle télique du nom auquel il est associé.

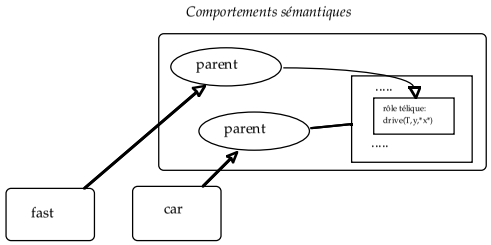


Figure 4.2 : "Fast car"

Une représentation qui structure les informations de manière hiérarchique est toujours difficile à modifier : il est coûteux voire impossible de modifier des liens d'héritage sans remettre en cause l'intégralité du domaine représenté. A l'inverse, une approche prototypique est plus ouverte pour une représentation évolutive des liens entre les éléments représentés.

[3] La notion de délégation sera exposée infra.

On peut ainsi envisager différentes solutions pour illustrer une approche à la Pustejovsky à l'aide de prototypes qui associerait aux entrées lexicales des scénarios guidant l'interprétation.

- Solution n°1 : on crée des hiérarchies locales.

On crée une structure définitoire de base qu'il convient d'affiner si nécessaire pour produire de nouvelles instances prototypiques.

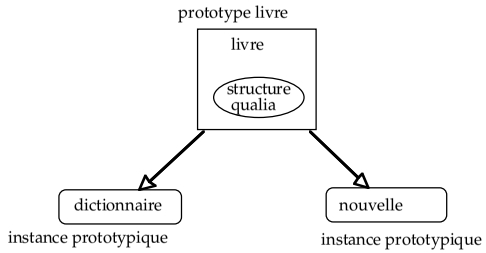


Figure 4.3 : Des prototypes comme moules définitoires pour les entrées lexicales.

Les instances prototypiques de `"dictionnaire"` et de `"nouvelle"` utilisent la définition du prototype `"livre"` comme s'il s'agissait d'un moule définitoire (à la manière d'une approche `Objet(Instance de Classe)`). Chacune d'elles porte une structure définitoire adaptée à l'entité représentée.

- Solution n°2 : toutes les entités sont au même niveau.

On considère un élément comme le plus typique d'une famille d'éléments à représenter. On peut ainsi amorcer la représentation en utilisant cet élément initial.

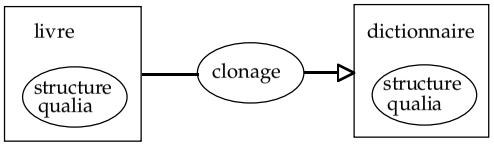


Figure 4.4 : De la multiplication des entités scénarisées.

L'opération de clonage consiste en une copie du prototype initial, on peut ensuite ajuster les éléments définitoires contenus dans la structure prototypique résultante pour l'adapter à la nouvelle entité représentée. Cette solution peut se combiner avec la mise en place de hiérarchie de comportements partagés par les entités représentées. Si le rôle télique est caractéristique des entités représentées, on peut factoriser les autres éléments définitoires.

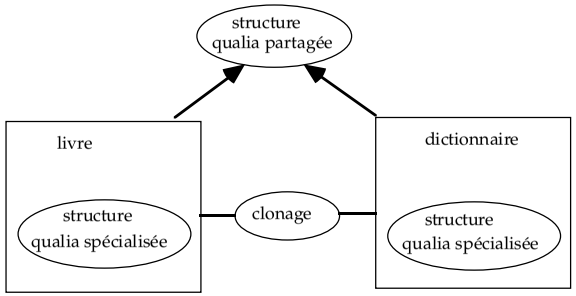


Figure 4.5 : Des scénarios adaptés à la réalité prototypique.

Par défaut, le comportement sémantique initial est composé du comportement sémantique défini par la structure qualia partagée et par la structure qualia spécialisée caractéristique de l'entité représentée. On peut ajuster cette dernière structure (par spécialisation), on peut aussi être amené à modifier la délégation de comportement partagé si nécessaire, ou à la spécialiser elle aussi (en ordonnant la délégation).

- Solution n°3 : tous les éléments sont définis en termes de prototypes.

On ne distingue plus la notion de comportement de la notion de prototype. Les entrées lexicales sont associées à des structures prototypiques qui peuvent être décrites par une association (à définir) avec des prototypes qui précisent leur rôle sémantique. Les entités voisines peuvent se définir par des opérations de copie et d'ajustement des valeurs associées.

éléments prototypiques

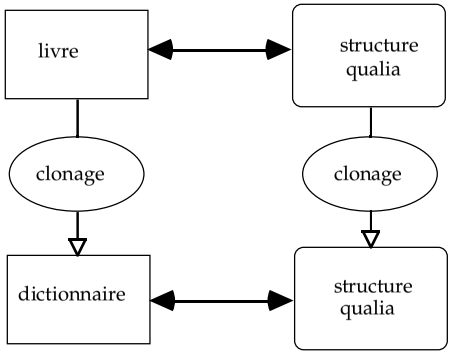


Figure 4.6 : Tout est "prototypisable".

- Solution n°4 : on ordonne la représentation prototypique précédente.

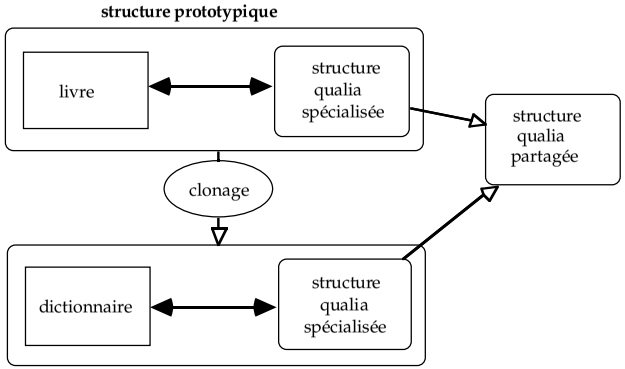


Figure 4.7 : Mise en ordre et prototypie.

Comme précédemment, on hiérarchise le savoir sur le domaine représenté en factorisant les comportements (en particulier les comportements sémantiques).

Il convient une nouvelle fois d'insister sur la nécessité de construire des mécanismes qui proposent des alternatives pour la construction du sens : il s'agit de définir les chemins de sens potentiels pour une entrée lexicale puis d'organiser au mieux les directions de ces parcours sémantiques afin de répondre à des besoins particuliers que les configurations interprétatives rencontrées révèleront. Ces comportements sémantiques doivent traduire les phénomènes de reprise, d'extension, de précision qui opèrent dans le processus d'interprétation. Le choix d'une telle approche vise donc à opérer une réflexion sur l'organisation des savoirs sémantiques que l'on peut attacher à une entrée lexicale (ces savoirs étant assimilés à des comportements). En particulier, il ne faut pas perdre de vue la nécessité de structurer l'ensemble des comportements sémantiques définis. Une prototypisation des savoirs peut conduire rapidement à une dilution incontrôlable et donc peu efficace pour établir à terme la mise en ordre des éléments représentés. Cette mise en ordre doit malgré tout préserver la possibilité d'une mise à jour éventuelle des comportements initialement définis. Si le comportement sémantique attaché à une entrée lexicale peut être affiné, il faut pouvoir spécifier dans la représentation sémantique le détour sémantique rencontré parmi les chemins de sens existants.

Cela revient à traiter parallèlement les points suivants :

- Structurer le savoir sémantique pour chaque entité représentée tout en laissant ouverte la représentation de ce savoir.

- Organiser les savoirs sémantiques sur l'ensemble des entités représentées en définissant les moyens pour ajuster ce classement.

Pour illustrer le premier point, on peut d'ailleurs s'interroger sur les modalités à mettre en oeuvre pour bien cerner les couples discriminatoires qui permettent l'activation de processus sémantiques particuliers. Pustejovsky utilise des oppositions en termes de traits différentiels pour distinguer les comportements sémantiques qu'il est possible d'attacher à une entrée lexicale. Le problème se pose donc de savoir s'il peut y avoir recouvrement entre les couples discriminants et les comportements induits.

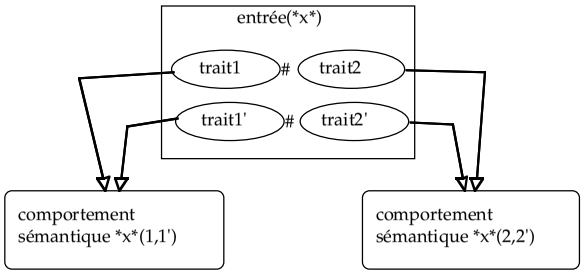


Figure 4.8 : Recouvrement de comportements sémantiques.

Sur quel type de "qualité" sémantique faut-il distinguer les couples discriminatoires activant des processus interprétatifs? Est-il (im)possible, souhaitable d'unifier ces types discriminatoires?

Dans la figure précédente, si on ne maintient pas la double opposition dans la définition des comportements sémantiques induits par ces traits, on risque de ne pouvoir activer le comportement sémantique adéquat. En effet si la structure définitoire ne tient compte que du couple (trait1,trait2), il peut s'avérer que ce couple soit non pertinent dans un contexte donné. Alors que le couple (trait1', trait2') le serait dans ce même contexte.

On rejoint ici encore des problèmes déjà mentionnés. Si ces oppositions constituent des savoirs établis, il faut les organiser pour ne pas laisser de côté des configurations interprétatives. Si ces savoirs sont révélés par l'analyse, il convient de les récupérer et de réorganiser les comportements sémantiques attachés aux objets concernés.

4.4.2.2. De l'unité dans la représentation

Une représentation hiérarchique structure les liens entre les noeuds inscrits dans le graphe d'héritage induit par cette représentation. La définition de mécanismes d'inférences lexicales dans une telle approche peut donc être assimilée à des parcours de ces liens d'héritage afin d'y saisir les informations pertinentes. Seuls les liens pré-définis sont valides; les inférences lexicales sont donc restreintes/contraintes par cette pré-définition de la hiérarchie initiale. Si une structuration du lexique ou des concepts sous forme d'une représentation hiérarchique permet de réaliser des liens d'héritage entre les éléments de la hiérarchie décrite, une telle structuration limite les liens entre les éléments représentés aux chemins qu'il est possible de parcourir dans le graphe d'héritage induit : la spécialisation traduit le passage de propriétés générales vers des propriétés particulières. Qu'en est-il des relations transversales qu'il peut être possible d'envisager entre les éléments d'un graphe d'héritage sur un domaine lexical ou conceptuel?

Le choix d'une représentation prototypique, c'est-à-dire le choix de l'absence, à priori, de structuration globale pour les éléments représentés, se heurte au problème crucial des liens à établir entre ces éléments surtout quand il s'agit de représenter des éléments qui agissent ensemble dans des processus de construction du sens. Il est donc impératif de définir les connexions qui existent entre les éléments représentés de manière prototypique : la prototypie doit s'organiser pour ne pas diluer les savoirs sur des entités qui s'ignorent. S'il s'agit de donner une représentation matérielle de la construction du sens en termes de prototypes, il faut que les éléments prototypiques définis aient les moyens de se (re)connaître mutuellement : enregistrer les liens initiaux entre les entités représentés et aussi pouvoir en construire.

De plus, il est important de bien distinguer ce qui relève de la représentation morphologique ou sémantique de ce qui relève de la représentation des processus qui manipulent ces précédentes formes de savoir. A priori puisque nous travaillerons dans un système qui ne manipule que des prototypes, ces derniers pourront être aussi bien des récepteurs de savoirs (figés ou non) que des passeurs agissant sur les précédents.

Il peut être délicat de restreindre les modalités de représentation de phénomènes "qualitativement" distincts sur un type particulier d'éléments. S'il peut s'avérer adéquat d'utiliser des prototypes pour décrire des problèmes précis, ce mode de représentation peut se trouver en difficulté pour la représentation d'autres domaines de savoir. Il ne s'agit donc pas de dire que le choix des prototypes s'impose pour traiter globalement le problème de la représentation matérielle des faits de langue, mais qu'il peut être adapté pour aborder certains phénomènes à l'oeuvre dans les faits de langue. Là encore il ne s'agit pas de privilégier une pratique par rapport une autre : l'approche prototypique a, semble-t-il, beaucoup de mal à rendre compte de certains phénomènes relevant d'une organisation hiérarchisée et figée du savoir. Il faut entendre la difficulté précédemment énoncée en termes d'inadaptation pour traiter le type de problème visé et non en termes d'impossibilité.

4.5. Approche pour une formalisation prototypique de la construction progressive du sens des formes moulin à N2

4.5.1. Une première approche guidée par les hypothèses sémantiques développées par (Habert & Fabre 1993)

4.5.1.1. Un calcul sémantique à base de traits

Cette approche pour la construction du sens des formes moulin à N2 (qui est développée en détail infra (-> partie 2 chapitres 5, 6)) peut être résumée de la manière suivante :

- **Etape 1 :**
  Construction de la représentation du Problème.
  Approche pour une représentation par prototypes des éléments lexicaux.
  Mise en place d'une analyse automatique guidée par des prototypes.

- **Etape 2 :**
  Affinement progressif des mécanismes de représentation et d'analyse.
  Mise en place d'un traçage de l'analyse.
  Prise en compte d'un modèle sémantique pour la construction du sens des moulin à N2 : assimilation des résultats produits à des classes connues pour lesquelles on possède des critères d'évaluation et de reconnaissance. Avec notamment un ajustement dynamique de la sémantique attachée à moulin qui devient +humain dans la série des moulins à paroles. On donne infra une présentation des filtres qui permettent une classification provisoire des séquences étudiées.

- **Etape 3 :**
  Activation de procédures de contrôle qui permettent de modifier les données initiales si l'Etape 2 ne réussit pas :
  - modification des valeurs sémantiques de manière externe (à "la main" par une intervention de l'utilisateur).
  - ajout de filtres (pour les N2 déverbaux).

- **Etape 4 :**
  Mémorisation des "évènements critiques" dans l'Etape 3 qui ont permis d'atteindre à la reconnaissance d'une classe.
  Ces évènements amènent à repenser le modèle sémantique choisi pour prendre en compte les faiblesses révélées par l'analyse.

Le calcul sémantique mis en place dans les Etapes 2 à 4 peut être décrit comme suit :

Soit Traits l'ensemble ordonné des traits sémantiques utilisés par le calcul sémantique :
Traits = {abstrait, concret, nParole, ali, synth, energie, comptable}

Soit Mots l'ensemble des mots x du lexique tels que les formes moulin à x existent.

Soit $Trait_t$:

Mots -> { +,-}
mot -> $Trait_t$(mot), valeur du trait sémantique t pour le mot considéré

Soit Sem :

Mots -> {(x1,...xi,...)/ xi= +,- et i=2....}
mot -> ($Trait_{t1}$(mot),$Trait_{t2}$(mot),....,$Trait_{tn}$(mot))

où n est le nombre de traits sémantiques utilisés par le calcul sémantique des formes moulin à N2 et où les ti appartiennent à Traits.

Soit Classes l'ensemble des classes sémantiques connues :
Classes = { PourSortie1, PourSortie2, PourSortie3, PourSortie4, PourSortie5, PourSortie6, PourEntrée1, PourEntrée2, AvecEnergie, AvecPartieDe}

Chaque élément de Classes correspond à un filtre de traits sémantiques associé à une sémantique possible pour le composé. Ces filtres sont des suites ordonnées de valeurs des traits sémantiques (abstrait, concret, nParole, ali, synth, energie, comptable) définies comme suit :

PourSortie1 = (-,+,+,-,-,-,+)
PourSortie2 = (-,+,+,-,-,-,-)
PourSortie3 = (-,+,-,+,+,-,+)
PourSortie4 = (-,+,-,+,+,-,+)
PourSortie5 = (-,+,-,-,+,-,-)
PourSortie6 = (-,+,-,-,+,-,+)
PourEntrée1 = (-,+,-,+,-,-,-)
PourEntrée2 = (-,+,-,-,-,-,-)
AvecPartieDe = (-,+,-,-,-,-,+)
AvecEnergie = (-,+,-,-,-,+,-)

Soit x ∈ Mots tels que moulin à x existe.

Calcul sémantique de moulin à x

1 Si Sem(x) ∈ Classes, l'analyse réussit.

2 Si Sem(x) ∉ Classes :
Dans ce cas, on est confronté aux alternatives suivantes :

2.1. Ajout de Sem(x) à Classes :
Classes:= Classes U {Sem(x)}. Il convient donc dans ce cas de réajuster le calcul sémantique pour prendre en compte la nouvelle classe reconnue.

2.2. Modification des valeurs sémantiques de x :
Pour t ∈ Traits, $Trait_t$(x):= + ou - . Dans ce cas, il s'agit plutôt de modifier les valeurs sémantiques initiales attachées au N2 pour "faire entrer" le composé produit dans le modèle construit. Après une mise à jour des valeurs sémantiques, on opère une nouvelle analyse (retour supra Calcul sémantique).

2.3. On ne fait rien, mais on garde en mémoire la configuration sémantique qui a échoué. Dans ce cas, il s'agit ni de forcer la représentation (2.2) ni d'intégrer trop tôt une nouvelle classe (2.1). On en reste au modèle initial tout en gardant en mémoire les faiblesses de celui-ci.

2.4. On évalue d'autres traits sémantiques ou morphologiques non encore utilisés par le calcul sémantique initialement défini. On garde aussi en mémoire la configuration sémantique qui a permis d'atteindre ou non à la reconnaissance. Il s'agit en fait d'une extension du modèle sémantique initial qui agit notamment pour le traitement des N2 déverbaux.

4.5.1.2. Limites de ce type de représentation sémantique en termes de traits

Il semble en effet délicat de restreindre la définition des représentations sémantiques en termes de traits binaires. Tout d'abord, on est souvent confronté dans la description des unités linguistiques à des données qui ne se laissent pas appréhender par une simple opposition entre deux possibles. Il s'agit au contraire d'évaluer des propriétés qualitatives dont la représentation ne se prête pas à une telle réduction. Il convient de réfléchir à des descriptions plus raffinées pour appréhender les entités représentées. Ensuite, la prise en compte de nouvelles informations peut conduire à une multiplication incontrôlable du nombre de ces traits binaires. De plus est-il légitime de définir une unité de langue comme une liste d'étiquettes qu'on empile? Qu'est-ce qu'un trait en général? Peut-on prendre n'importe quoi? Que recouvrent réellement ces attributs? (Sur ces problèmes (Rastier & al. 1994)). Dans la mesure où le travail développé ici vise à représenter la construction du sens au travers de processus interprétatifs qui agissent à partir d'une signification minimale que les informations rencontrées sur une configuration donnée permettent de préciser, il ne semble pas adéquat de figer les informations sous une forme de liste d'étiquettes disparates. Il s'agit au contraire de préciser les chemins de sens que l'on peut construire sur un item donné (reconstruire un espace sémantique possible en termes d'opérations sémantiques) et dont les modalités de descriptions finales (sous forme de traits dont la nature sera à préciser, ou sous un autre forme) seront déterminées et ajustées suivant les directions de sens mises en place pour cet item.

4.5.2. `Moulin` : (Re)construire une Représentation Sémantique Prototypique[4]

Nous considérons tout d'abord la forme initiale de description de `moulin` adoptée supra :

`moulin` : nom masculin singulier

Si l'on considère `moulin` comme un représentant typique des mots généralement regroupés dans la catégorie grammaticale des noms, on peut considérer qu'il peut servir de modèle pour la création d'un prototype `nom` (noté prototype `N`). Cette description sommaire que l'on peut faire de ce mot va permettre la création du prototype `N` qui servira de base pour la création de nouveaux prototypes `N`.

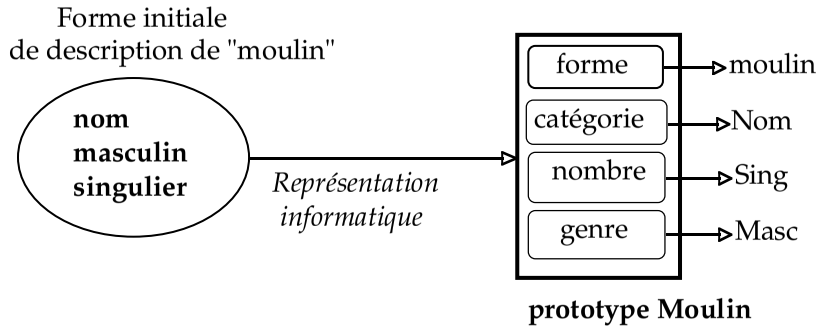


Figure 4.9 (1) : `Moulin` est à l'origine de la création du monde des noms : première étape, la création.

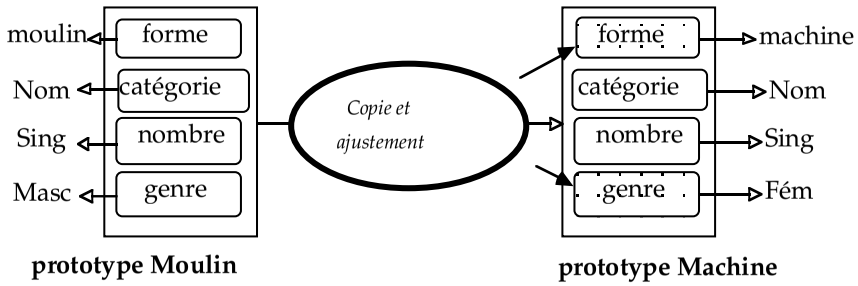


Figure 4.10 (2) : `Moulin` est à l'origine de la création du monde des noms : deuxième étape, la reproduction.

Sur notre exemple, `machine` est un autre représentant de la catégorie grammaticale des noms, il possède donc au moins les mêmes attributs définitoires que `moulin`. On peut donc créer un prototype associé à `machine` à partir du prototype initial associé à `moulin`. Les seuls attributs à modifier pour représenter ce nouveau mot sont les attributs décrivant la forme et le genre de celui-ci. Essayons maintenant d'attacher une représentation sémantique à `moulin`. Il

[4] On donne infra une présentation de l'implémentation proposée pour traiter les problèmes décrits ici (-> partie 2 chapitres 5, 6 -> annexe partie 2 chapitre 5).

s'agit ici de définir l'information sémantique en termes de comportements particuliers définis par les configurations interprétatives rencontrées. Pour cela, nous souhaitons attacher à `moulin` une représentation sémantique initialement peu déterminée et ensuite affiner cette représentation pour tenir compte des extensions ou des changements de sens qu'il est possible de rencontrer dans certaines réalisations contextuelles.

4.5.2.1. `Moulin` emballe la machine prototypique

Notre prototype initial attaché à `moulin` peut donner naissance à deux représentants prototypiques distincts dans leurs comportements sémantiques. Cette première décomposition ne dit encore rien de la description sémantique qui est attachée à chacun des prototypes créés. Elle met malgré tout en lumière les problèmes à résoudre pour construire une représentation des connaissances de manière prototypique.

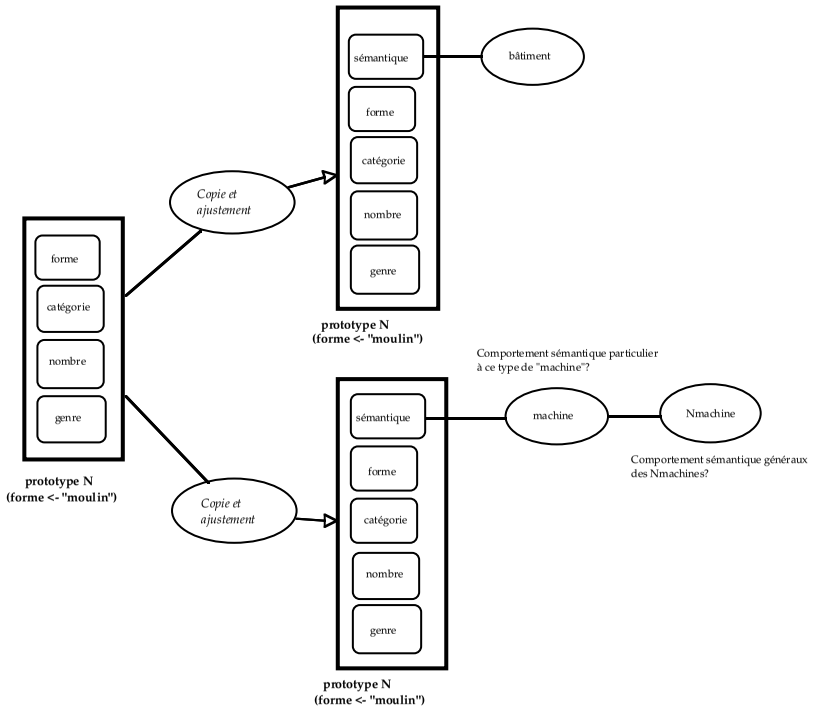


Figure 4.11 : `Moulin`, quelle représentation sémantique?

Tout d'abord, si on considère que `moulin` est effectivement un représentant typique des noms de machines, il peut servir de base pour la création d'autres exemplaires de cette classe de noms. Il reste évidemment à vérifier si cette classe de noms organise des comportements sémantiques particuliers, et parmi ces comportements, quels sont ceux qui intéressent `moulin` ? Ensuite, cette décomposition amène à dresser la liste de tous les possibles de sens de `moulin`. Pour reprendre Pustejowsky (Pustejowsky 1991), on peut essayer au contraire de capter les régularités lexicales et sémantiques puis spécifier les réalisations dans la représentation.

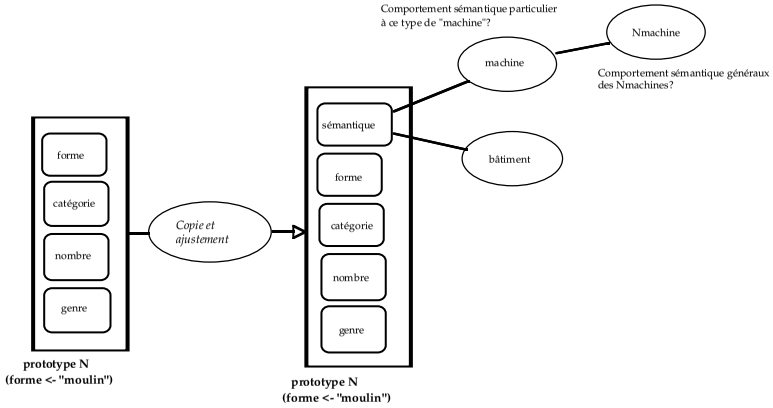


Figure 4.12 : Un prototype `moulin` qui capte les régularités lexicales et sémantiques.

Pour aller plus loin, on peut aussi vouloir définir sur moulin un noyau de sens ("à la Fuchs") qui organise la construction sémantique du prototype en fonction des informations auxquelles celui-ci aura accès dans un contexte donné. Le noyau de sens que nous visons à définir se place dans une perspective d'ébauche de sens. Il s'agit de ne pas émietter dans la représentation des sens possibles, mais au contraire de partir d'une représentation "grossière" puis de la modifier si nécessaire.

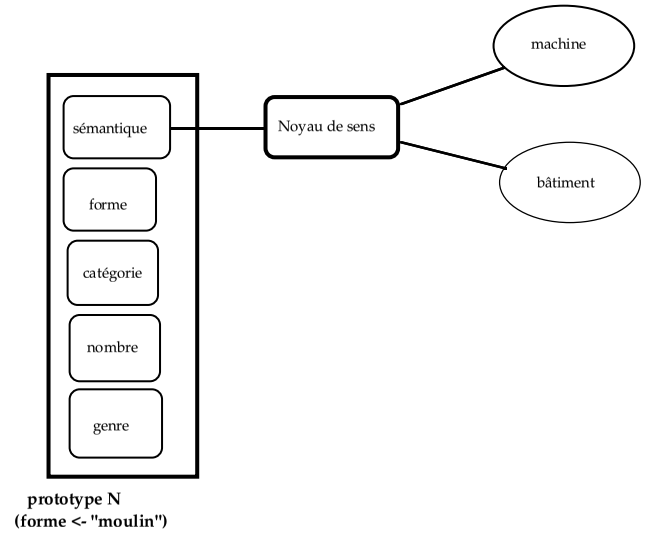


Figure 4.13 : `Moulin`, un noyau de sens ?

Là encore, on est confronté à des choix importants. Si on s'oriente vers une représentation d'un noyau de sens qui va du "général au particulier" sémantique, il convient de ne pas perdre de vue que les descriptions faites sur `moulin` peuvent intéresser d'autres représentants lexicaux "sémantiquement proches". Sans présumer de ce que sera la description sémantique de `moulin`, on peut par exemple vouloir décrire précisément les comportements sémantiques des noms d'instruments et se servir de cette description particulière de `moulin` pour amorcer une

description sémantique des noms d'instruments qu'il conviendra de préciser par une étude complémentaire d'autres noms d'instruments particuliers. On peut ainsi considérer que la transformation `machine -> lieu`, non articulée dans la représentation précédente, est une opération générale qu'il est possible d'appliquer à bon nombre de noms de machine. Cette transformation dépendant en partie de la taille de la machine visée : `"Je suis entré dans le réacteur"` vs `"Je suis entré dans le concasseur"`. Nous revenons donc sur la représentation initialement adoptée afin d'articuler cette transformation de sens qui peut s'appliquer sur le noyau de sens associé à `moulin`.

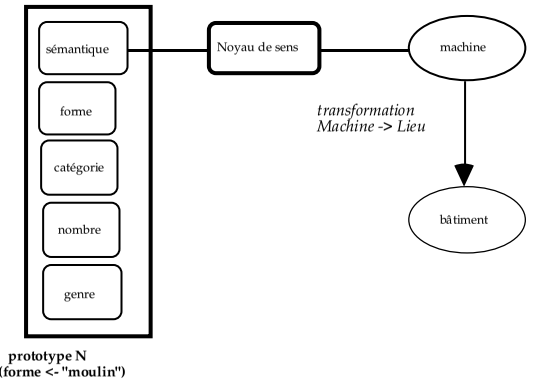


Figure 4.14 : `Moulin`, un noyau de sens qui articule les transformations de sens observées

Le noyau de sens attaché à `moulin` doit tendre à tracer une direction "sémantique" adéquate pour la description de `moulin` comme `instrument`. Il s'agira ensuite de détacher dans les comportements sémantiques définis ceux qui relèvent plus généralement de la classe des noms d'instruments, tout en permettant de retrouver ces informations particulières dans cette classe de noms et notamment celles qui concernent le comportement sémantique de `moulin` (mise en oeuvre d'héritage de comportement sémantique). On peut d'ailleurs se demander si ce processus d'abstraction est cohérent avec une approche prototypique.

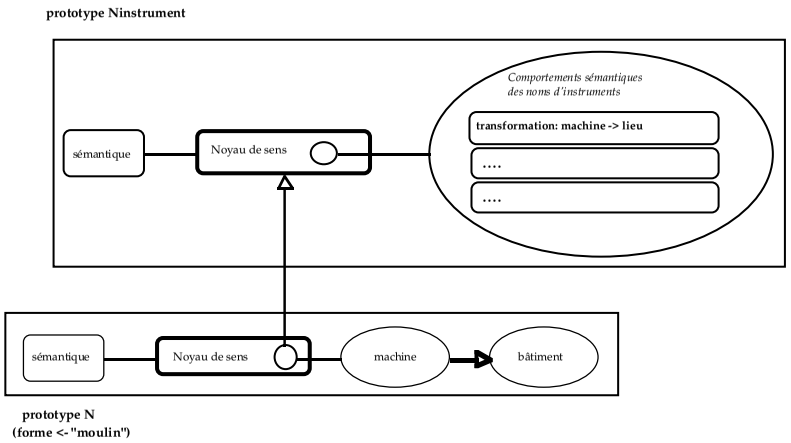


Figure 4.15 : Certaines extensions peuvent s'appliquer aux noms d'instruments.

4.5.2.2. Des `moulins` aux `moulins à N2`

Il est possible à ce stade d'étendre la représentation sémantique de `moulin`. Il peut être associé à un prédicat dont la valeur prédicative dépend du `N2`. Si `N2` est un produit naturel, la valeur prédicative traduit une transformation. Si par contre, le `N2` est un artefact, celle-ci traduit plutôt une création. On se place donc ici dans une approche qui élargit la notion de prédicativité nominale à des constituants autres que des déverbaux. Ces nouvelles informations nous permettent de préciser notre représentation prototypique de la manière suivante[5].

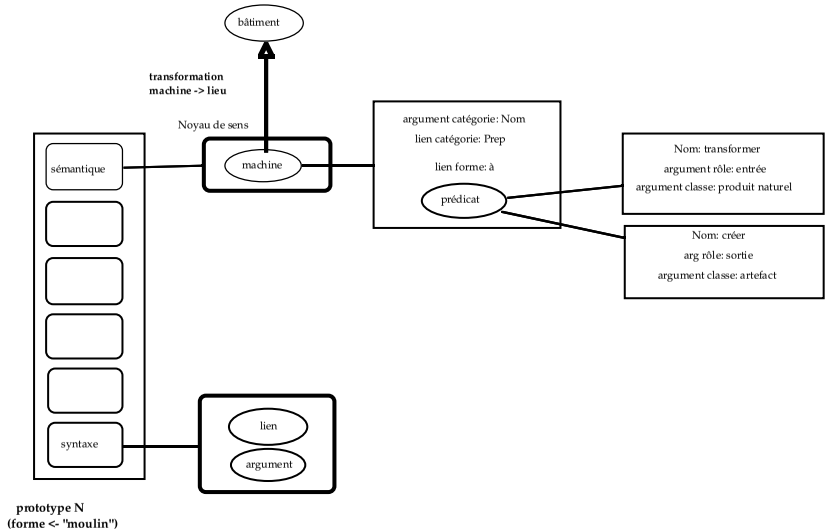


Figure 4.16 : Des `moulins` aux `moulins à N2`.

Les réalisations telles que `moulin à vent`, `moulin à eau`, `moulin à bras` sont partiellement prises en compte dans cette représentation sémantique attachée à `moulin` : la valeur prédicative `transformer` reste valide, par contre la classe de l'argument doit être étendue à des arguments de type `énergie`. De plus la nature de la relation est elle aussi affectée, il s'agit plutôt d'une relation `partie-de`. L'argument est dans une relation de partie constitutive dans le processus de transformation : le moulin transforme une énergie (naturelle en général) en énergie mécanique ou autre, et l'énergie (en entrée) est partie intégrante du processus de transformation réalisé par le moulin. Il semble d'ailleurs que cette dernière remarque nous permette de préciser davantage la représentation prédicative associée à `moulin` . On peut en effet considérer que `moulin` implique un changement d'état qui recouvre plusieurs réalisations lexicales suivant que l'argument soit un objet fabriqué, un produit naturel ou un nom d'énergie. On peut ainsi définir un prédicat générique unique `faire-changer-d'état` dont les manifestations lexicales dépendent de l'argument : cette solution est assez proche de la co-composition de Pustejowsky.

[5] La représentation proposée ici ne prend pas encore en compte des séquences telles que "un moulin à blé à vent" ou "*un moulin à farine à blé"; ce qui revient à distinguer les status de l'argument et du modifieur, on y reviendra par la suite.

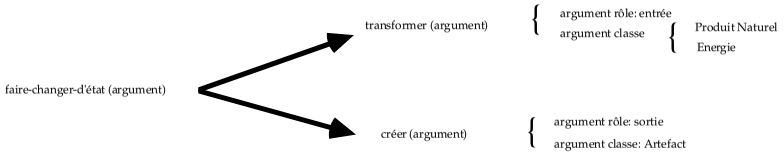


Figure 4.17 : Un prédicat générique pour `moulin`.

Ce choix permet tout d'abord de maintenir une représentation qui s'accorde avec bon nombre de noms d'instrument; ensuite, la sous-spécification du prédicat en `transformer` ou `créer` peut être vue soit comme un problème de spécialisation lexicale, soit comme un phénomène de co-compositionalité. La figure 4.18 reprend la représentation précédente en ajustant la définition de la nature du prédicat associé à `moulin`.

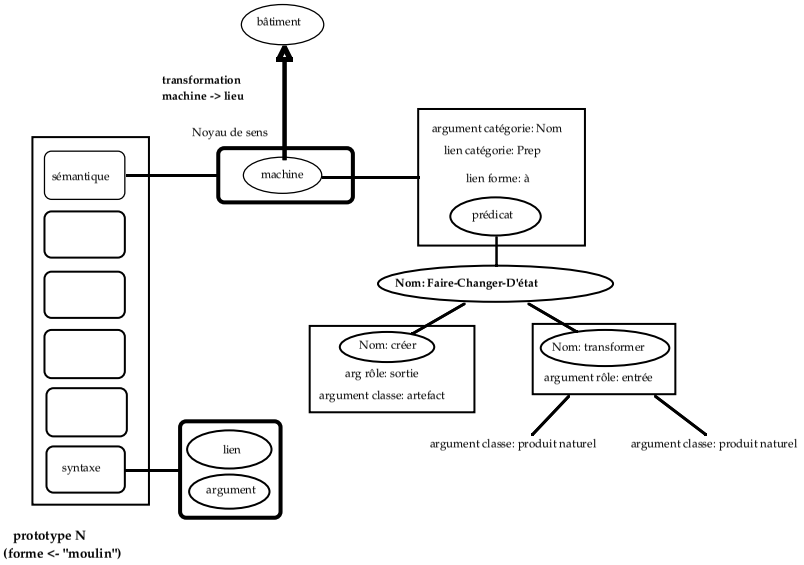


Figure 4.18 : Une représentation prédicative sous-spécifiée pour `moulin`.

En ce qui concerne les emplois comme `moulin à bras`, `moulin à cylindres`, `moulin à meules`, les arguments décrivent des éléments intégrés au moulin (l'argument est dans une relation de partie constitutive de l'instrument décrit) : cette caractérisation par la facette `partie-de` met d'ailleurs en évidence des parties saillantes du moulin décrit. On peut définir un rôle constitutif (à la Pustejovsky) ou un rôle aspectal (à la Fradin) pour prendre en compte ce genre d'extension. La valeur prédicative relativement neutre peut d'ailleurs rester de mise tant que l'on n'en sait pas plus sur le moulin visé : quitte à spécifier et/ou sous-spécifier cette valeur si nécessaire (-> figure infra).

4.5.2.3. Des moulins à prières aux moulins à paroles

`Moulin à prières` est semble-t-il central pour la prise en compte de certains sens figurés du patron `moulin à N2`. A priori deux attitudes auraient été possibles pour inscrire `moulin à prières` dans la représentation sémantique initiée ci-dessus. Si `moulin à prières` est,

dans la définition précitée, un instrument[6] constitué d'un "*cylindre renfermant des bandes de papier recouvertes de formules sacrées pour acquérir les mérites attachés à la répétition de cette formule*", on peut considérer que ce type d'instrument transforme des formes/signes graphiques en source sonore : les prières (i.e. il s'agit en quelque sorte d'une opération de décodage). Si l'on considère maintenant la relation qui existe entre `tourne-disque` et `disque`, on peut considérer que cet instrument opère un décodage du sillon inscrit sur le disque en une production de sons. A l'image de cette relation, on pourrait envisager d'associer à `moulin à prières` le prédicat `transformer` avec `[argument rôle : sortie]`, à la différence de la relation de transformation `tourne-disque` et `disque` pour laquelle `[argument rôle : entrée]`. Et c'est justement ce dernier point qui semble le plus inadéquat pour retenir le prédicat choisi. En effet, le processus de transformation privilégie fortement la donnée de l'objet à transformer plutôt que la donnée de l'objet issue de la transformation. Au niveau conceptuel, un tel choix serait donc maladroit pour décrire un instrument de transformation pour lequel l'objet transformé serait masqué. De plus, au niveau de la représentation, ce choix se serait révélé coûteux : un tel choix impose de disposer de filtres supplémentaires pour le prédicat choisi et surtout on verra que la solution suivante se révèle plus productive, notamment pour la prise en compte des sens figurés du patron `moulin à N2` tels `moulin à paroles` ou `moulin à conversation`. Si l'on choisit en effet une solution qui associe à `moulin à prières` le prédicat qui traduit la création de l'argument donné, le lien logique entre le processus décrit et l'argument donné syntaxiquement est cohérent : le processus de création privilégie fortement la donnée de l'objet créé. Il faut donc compléter la définition des filtres initialement attachés au prédicat dans notre représentation sémantique précédente pour prendre en compte la classe des `NParole` pour le prédicat `créer` (`[argument classe : NParole]`). Ensuite, il semble possible, à partir de cette interprétation, de retrouver par dérivation les sens figurés que l'on rencontre pour les séquences `moulin à N2`. Les emplois tels `moulin à paroles` peuvent eux-aussi être associés au prédicat `créer` dans la mesure où il s'agit dans ce cas de la création d'un excès de paroles[7]. `Moulin` désigne donc dans ce cas une entité de type `+humain`. Nous devons donc ajuster ici d'ajuster la représentation du noyau de sens : `moulin` "ajuste" certains traits de sa définition sémantique initiale de `machine`; il ne s'agit plus de décrire initialement une machine dont les caractéristiques sont précisées par la structure argumentale donnée, mais il s'agit de décrire une personne[8] qui telle une machine produit une grosse quantité de paroles (c'est aussi la structure argumentale qui permet cette dérivation de sens). En termes de représentation prototypique, si l'on considère que `moulin` est typiquement une machine, on doit malgré tout pouvoir retrouver ce détour par rapport à la préférence sémantique attachée à `moulin` si les informations rencontrées le nécessitent. Cette nouvelle direction de sens pour `moulin` peut se réaliser sous la forme d'un clonage du prototype initial attaché à `moulin` (le prédicat est maintenu : `créer, argument rôle : NParole, argument : +excès`) avec parallèlement une assignation d'un nouveau type de comportement sémantique qui décrit l'attachement de ce représentant de `moulin` à la classe des `Nhumains`. On peut d'ailleurs remarquer que l'interprétation initiale de `moulin à prières` contient déjà le trait `excès` dans la production décrite par l'argument et que ce trait est maintenu dans le prototype dérivé. De plus, cette solution permet de concilier une autre interprétation pour `moulin à prières` : il peut s'agir d'une autre dénomination pour `grenouille de bénitier` (Habert & Fabre 1993). La figure 4.19 tente de rassembler les points de modification annoncés dans la Représentation Sémantique Prototypique attachée à `moulin`.

[6] La solution 264 du le corpus donné en annexe en donne une définition précise: "*// un thibétain inscrit sa prière sur une bande et l'enferme dans un cylindre qu'il fait tourner dans un moulin à prières...*"

[7] Entendu dans la bouche d'un locuteur canadien de langue française: "Les français sont des verbaux moteurs... de vrais moteurs à parler".

[8] La solution 12 du corpus présente également la forme "moulin à papier" auquelle on peut associer "personne qui produit une grande quantité de papier" i.e. un copiste.

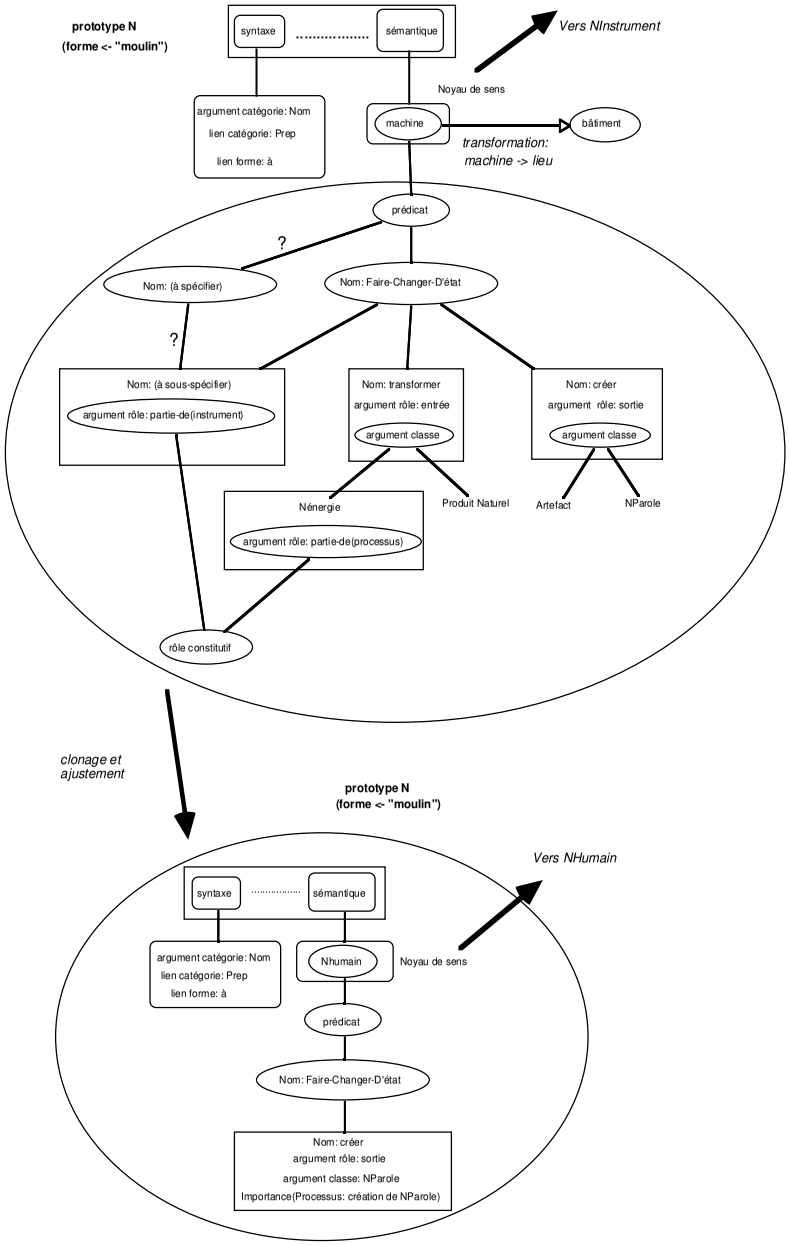


Figure 4.19 : Représentation Sémantique Prototypique.

Les emplois métaphoriques s'accordent semble-t-il avec la valeur prédicative retenue : il y a un changement d'état du processus de création en sortie de l'argument donné, celui-ci devient important et le résultat produit est massif.

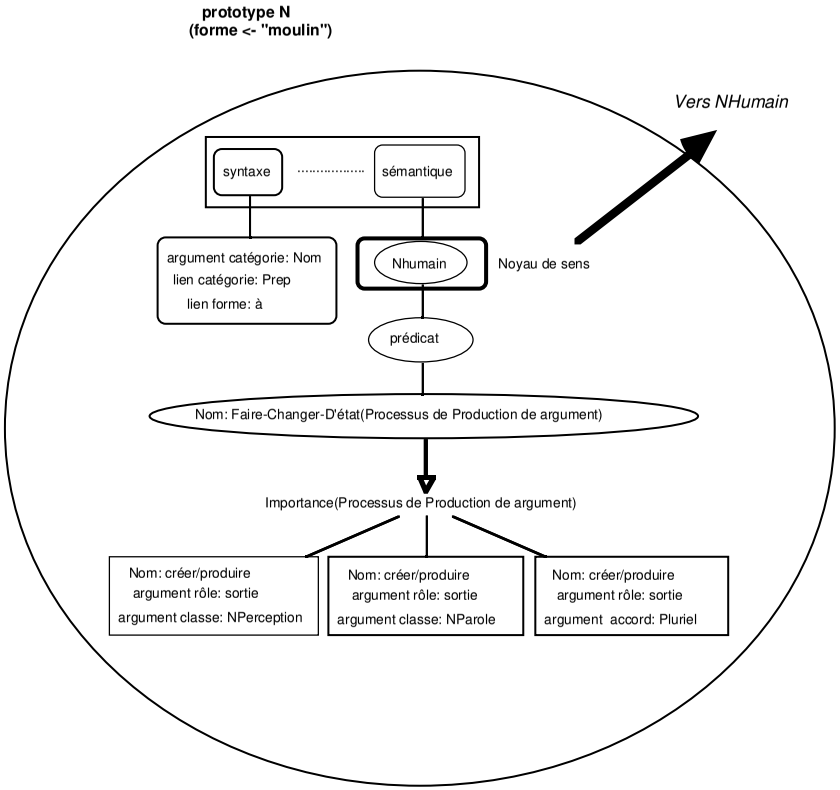
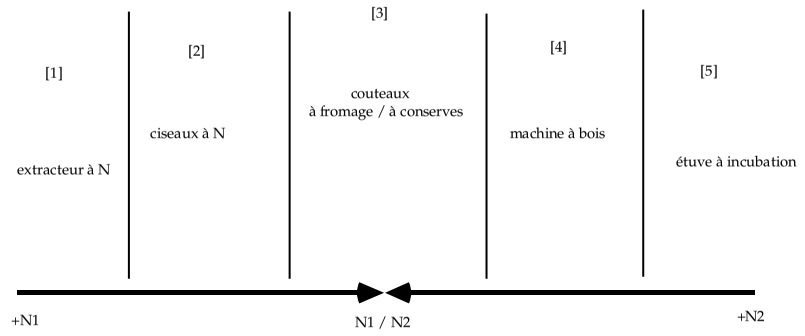


Figure 4.20 : Représentation Sémantique Prototypique pour `moulin Nhumain`.

4.5.2.4. `Moulin` et les déverbaux

Comme on l'a vu précédemment, certains exemplaires de `moulin à N2` ne décrivent pas ce que produit le moulin ni ce qu'il transforme, mais ils représentent un instrument constitué par le moulin. On retrouve là des processus généraux qui sont à l'oeuvre dans de nombreux composés en `N1 à N2` : une combinaison de deux prédicats qui s'articulent suivant la part que prend chacun des constituants du composé. On trouve par exemple dans (1) `étuve à incubation` ou (2) `ventilateur à aspiration` le même genre de relation entre un processus secondaire (le moyen) et un processus primaire (la fin) qui se réalise en sens inverse : (1) l'incubation se fait au moyen de l'étuve, (2) la ventilation se fait au moyen de l'aspiration. L'information sémantique portée par le `N2` peut ainsi prendre le pas sur celle du nom tête : dans (1) le `N2 déverbal` "pèse" sur la détermination du prédicat sous-jacent; alors que dans (2) c'est le "poids" du `N1 déverbal` qui importe. On rejoint donc ici une analyse des composés en `N1 à N2` qui force à réévaluer la place du `N2` dans la combinaison et qui tend à proposer un mode de calcul du sens en fonction de la nature des deux `N`. Le travail initié par (Habert & Fabre 93) s'oriente d'ailleurs depuis vers une analyse des `N à Ni` (noms composés désignant des

instruments) en suivant le principe de co-compositionalité. Les différentes analyses possibles s'ordonnent en effet sur un continuum en fonction de la part de chacun des composés donnés en suivant une échelle de co-compositionalité. On reprend ici le schéma proposé par Cécile Fabre :



[1] = N1 déverbal [3] = N1 prédicat sous-spécifié [5] = N2 déverbal
[2] = N1 nom rôle [4] = N1 hyperonymique

Figure 4.21 : La co-compositionalité pour l'analyse des `N à Ni`.

Dans les réalisations `moulin à (N2 déverbal)` mentionné supra, il en va ainsi pour la détermination du prédicat sous jacent : la ventilation se fait au moyen du moulin. Que reste-t-il malgré tout de la valeur prédicative attachée à moulin? Il reste possible de dire que l'on a un changement d'état de la valeur prédicative initialement attachée à `moulin` i.e. `Faire-Changer-D'état` (le prédicat comme agent de la réflexivité!) : le prédicat s'applique à lui-même si le `N2` est un déverbal et change sa valeur prédicative.

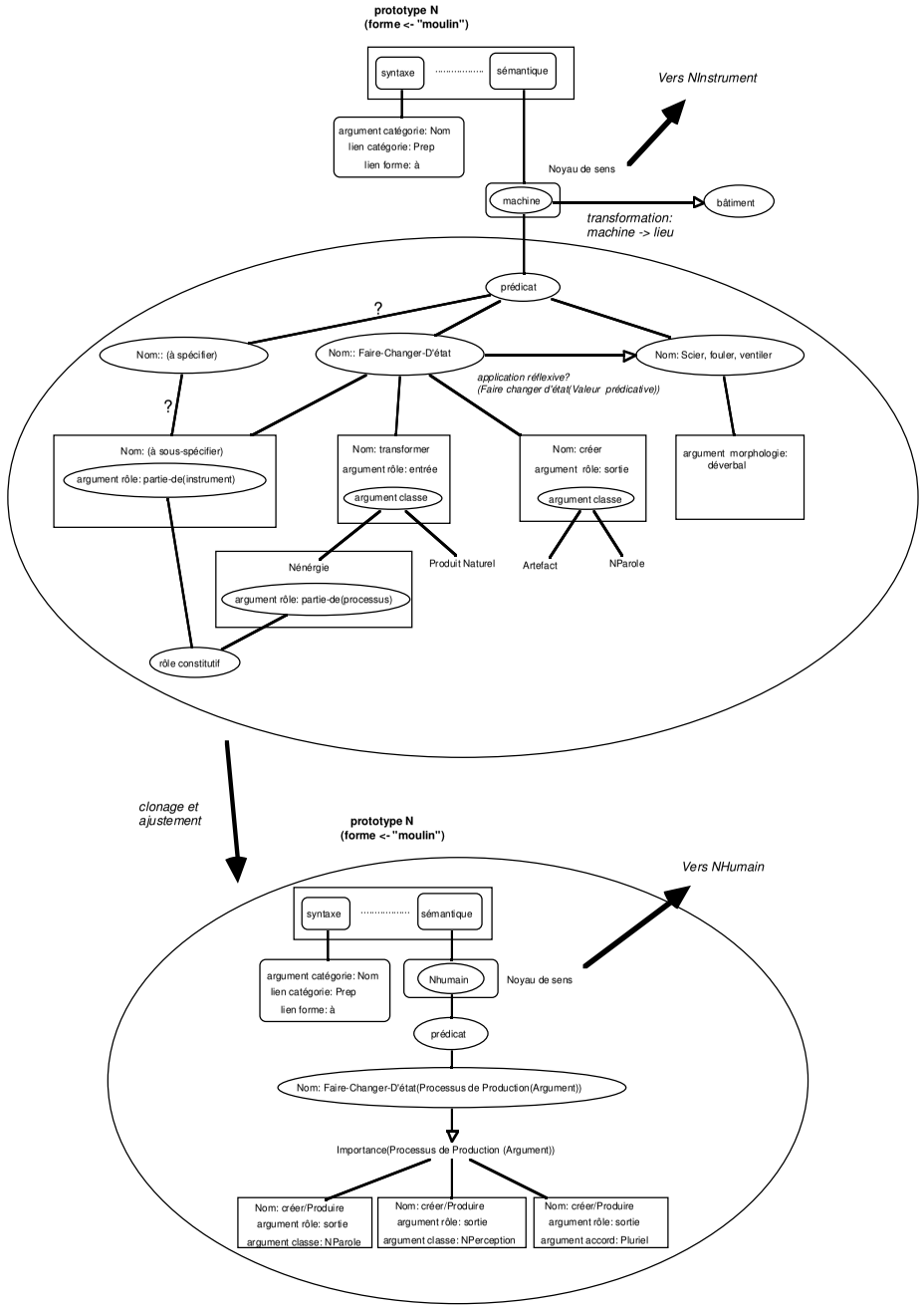


Figure 4.22 : Représentation Sémantique Prototypique affinée.

4.5.2.5. Articuler les niveaux de représentation

Cette représentation sémantique provisoire nous permet de repréciser la définition initiale attachée à `moulin` :

Un moulin[9] est :

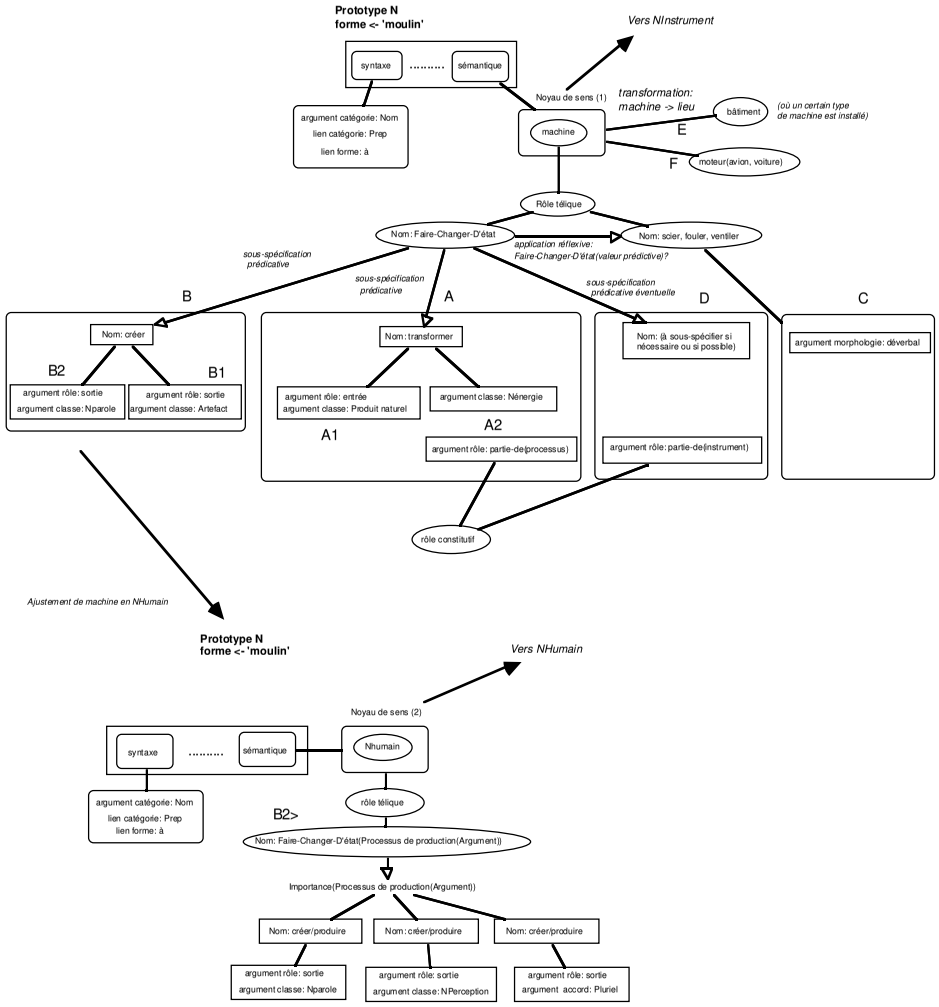
- une machine :
  - dont le type est précisé par la structure argumentale qui lui est associée :
    `(tête : moulin , lien[catégorie : Préposition, forme : à], argument[catégorie : Nom]),`
    `(prédicat : Faire-Changer-D'état)`
    - une machine dont l'argument donné est un composant du moulin (`argument rôle : partie-de`) ($moulin_D$) ; la valeur prédicative donnée peut être à sous-spécifier ou à respécifier si nécessaire i.e. si des informations contextuelles supplémentaires le permettent.
    - une machine qui transforme un produit naturel ($moulin_{A1}$) : la valeur prédicative est sous-spécifiée en `transformer, [(argument classe : produit naturel) , (argument rôle : entrée)]` ; elle peut être aussi une machine qui transforme une source d'énergie en une autre source d'énergie ($moulin_{A2}$) : la valeur prédicative est sous-spécifiée, de la même manière, en `transformer [(argument classe : Nénergie), (argument rôle : partie-de (processus))]`, ce dernier point, secondaire, indique simplement que la source d'énergie en entrée est partie intégrante du processus de transformation décrit.
    - une machine qui crée un artefact ($moulin_{B1}$) : la valeur prédicative est sous-spécifiée en `créer, (argument classe : artefact)` ; elle est aussi une machine qui produit des prières ($moulin_{B2}$) : la même valeur prédicative `créer` s'applique, `(argument classe : NParole)` ; ($moulin_{B2>}$) par extension décrit aussi un être humain qui telle une machine crée un excès de paroles, de conversation, (dans ce cas `(argument classe : NParole)` entraîne `(tête classe : Nhumain)`) ; la valeur prédicative sous-spécifiée traduit de plus le changement d'état du processus de production de l'argument donné : production/création importante ou massive.
    - une machine qui scie, ventile, foule et dont le moulin est un élément intégré à cette machine (`argument morphologie : déverbal`) ($moulin_C$) ; ici, c'est la valeur prédicative initiale qui a changé d'état.
  - du type moteur d'avion ou de voiture ($moulin_F$) ; on peut d'ailleurs remarquer que cette interprétation rejoint fortement ($moulin_C$).

- par extension un bâtiment où un certain type de moulin (décrit supra) est installé ou l'a été ($moulin_E$) (sauf pour `moulin à prières` et pour les emplois métaphoriques) ;

Les emplois métaphoriques de `moulin` posent semble-t-il un autre type de problème (qui dépasse d'ailleurs le cadre de notre travail). En effet, on peut considérer que les métaphores de moulin envahissent le langage au moment où d'autres machines plus perfectionnées sont sur le point de détrôner le prototype de machine que constituent les moulins. On peut d'ailleurs tracer un parallèle entre ce développement des emplois métaphoriques de moulin au moment du développement industriel et le développement des emplois métaphoriques de l'électricité que l'on rencontre aujourd'hui. L'électricité est elle aussi amenée à être supplantée par d'autres sources d'énergie. Ce passage vers d'autres sources d'énergie s'accompagne dans le même temps d'une multiplication là aussi d'emplois métaphoriques : *"Pourquoi, alors que le courant passait de moins en moins bien, on pétait les plombs de plus en plus chez ceux qui avaient deux fils qui*

[9] Les indices attachés aux définitions possibles pour `moulin` font référence aux "pistes de sens" que la figure qui suit présente

font masse. Fallait-il à ce point court-circuiter le langage officiel afin de créer un électrochoc dans la société? Manière, en quelque sorte, de se trouver quelques atomes crochus entres électrons libres. Histoire d'en finir avec les branchés et leurs passions cathodiques. Le langage électrique stigmatisait avant tout le dérèglement, le dérangement, la panne dans la continuité du circuit. Il vieillissait vite, mais son emploi en rajeunissait certains. Par contraste, il était remarquable de constater que l'informatique n'avait produit jusqu'alors qu'un vocabulaire, mais pas encore de langage, malgré les efforts de novlangue" (Gilles Tordjman in *Les Inrockuptibles* n°10 à propos du livre *Atlas occidental* de Daniele del Giudice). La principale difficulté que met en lumière cette nouvelle description sémantique de `moulin` est de faire appel à des savoirs de différentes natures pour fournir une interprétation possible ou des éléments pour un début d'interprétation. Le problème de l'articulation de ces savoirs ne se laisse pas appréhender sans difficulté. Les aspects sémantiques ne sont pas pertinents de manière régulière dans la construction des différentes interprétations possibles. La fonction, la constitution de l'objet décrit, la sémantique de l'argument etc. sont autant de facettes dont il faut tenir compte pour reconstruire une signification possible, d'autant plus que ces savoirs peuvent eux aussi agir en retour sur la représentation sémantique. Il reste malgré tout possible de tracer une esquisse des directions de sens qui semblent se dessiner dans cet ensemble d'interprétations. Si `moulin` est une machine, le type de machine en question sera explicité ou esquissé par les informations que l'on pourra récupérer dans la structure argumentale associée. La sémantique résultante n'est d'ailleurs pas figée : certaines informations sémantiques extérieures peuvent affiner le savoir attaché à `moulin`. Tous les moulins n'ont pas de rôle télique prédéfini dans cette représentation (ils peuvent malgré tout hériter d'une valeur prédicative générique comme dans les figures présentées supra) ; il n'est d'ailleurs pas sûr qu'ils en aient tous un; on peut cependant penser que `moulin à cylindres`, par exemple, peut recevoir des informations sémantiques supplémentaires dans un contexte donné qui précisent la représentation sémantique initiale que décrit la figure 4.23.



(1) traduit en particulier que moulin est un instrument concret, comptable, inanimé...
(2) traduit quant à lui que moulin est devenu +humain...

Figure 4.24 : `Moulin`, des pistes de sens.

Cette mise en ordre de l'interprétation ne limite pas les ajustements qu'il reste possible de faire : soit en précisant plus avant les aspects sémantiques décrits, soit en examinant de nouvelles séquences que les présentes hypothèses ne prennent pas encore en compte. Il reste en effet possible de dégrossir cette cartographie des directions de sens possibles en y greffant de nouveaux points d'entrée sur les noeuds existants ou en précisant les chemins déjà tracés.

Interprétations de `moulin` et représentations logiques :

En reprenant les notations de Fradin (Fradin 93) :

$Q_{tél}$ *"concerne les propriétés téliques au sens large : la fonctionnalité de l'objet aussi bien que son origine"* .
Q_{mod} *"concerne le mode d'existence de l'objet : essentiellement sa catégorisation par rapport aux enchaînements de causalité"* .
Q_{asp} *"concerne les propriétés aspectales et constitutives : rapports tout/partie, figure/site, et propriétés servant à identifier l'objet en question"* .

On peut construire les représentations sémantiques suivantes pour `moulin` :

- Interprétation A1 : `moulin à blé`
  `(moulin x)`
  $Q_{tél}$ = `(Transformer y x) et (ProduitNaturel y)`
  Q_{mod} = `(instrument x)`

- Interprétation A2 : `moulin à vent`
  `(moulin x)`
  $Q_{tél}$ = `(Transformer y x) et (Energie y)`
  Q_{mod} = `(instrument x)`

- Interprétation B1 : `moulin à farine`
  `(moulin x)`
  $Q_{tél}$ = `(Créer y x) et (Artefact y)`
  Q_{mod} = `(instrument x)`

- Interprétation B2 : `moulin à prières`
  `(moulin x)`
  $Q_{tél}$ = `(Créer y x) et (NParole y)`
  Q_{mod} = `(instrument x)`

- Interprétation B2> : `moulin à paroles`
  `(moulin x)`
  $Q_{tél}$ = `(Créer y x) et (Production(NParole) y) et (Important y)`
  Q_{mod} = `(NHumain x)`

- Interprétation C : `moulin à ventilateur`
  `(moulin x)`
  Q_{asp} = `(Partie y x) et (Instrument y)`
  $Q_{tél}$ = `(Faire-Changer-D'état)`
  Q_{mod} = `(PartieDe[Instrument] x)`

`(Partie x y)` se lit `(y a pour partie x)` i.e Instrument avec moulin.

- Interprétation D : `moulin à cylindres`
  `(moulin x)`
  Q_{asp} = `(Partie y x) et (PartieDe[Instrument] y)`
  $Q_{tél}$ = `(Faire-Changer-D'état)`
  Q_{mod} = `(instrument x)`

- Interprétation E :
  `(moulin x)`

$Q_{tél}$ = (Faire-Changer-D'état)
Q_{mod} = (instrument x)

- Interprétation F :
  (moulin x)
  Q_{mod} = (bâtiment x)

Une représentation sémantique pour `moulin` peut donner la structure suivante :

```
(moulin x)
Qasp1 = (Partie x z) et (PartieDe[Instrument] x)
Qasp2 = (Partie z x) et (PartieDe[Instrument] z)
Qtél0 = (Faire-Changer-D'état x y)
Qtél1 = (Faire-Changer-D'état x y) et (Artefact y)
Qtél2 = (Faire-Changer-D'état x y) et (ProduitNaturel y)
Qtél3 = (Faire-Changer-D'état x y) et (ProductionDeParoles y) et
(Importance y)
Qmod1 = (instrument x)
Qmod2 = (instrument z)
Qmod3 =(NHumain x)
Qmod4 =(Bâtiment x)
```

4.5.2.6. Préférences d'interprétation et prototypie

Si la notion de saillance est fondamentale dans une approche prototypique, une réflexion sur l'organisation des connaissances en termes de préférence d'interprétation reste à faire: on peut ainsi envisager de mettre en avant les interprétations "centrales". Toutes les interprétations n'ont pas le même "poids"; il semble donc cohérent d'articuler les savoirs représentés en tenant compte des réalités de production rencontrées.

Les résultats rassemblés ci-dessous permettent d'avancer que les séquences `moulin à N2` sont avant tout des instruments de transformation puis des instruments de création (aux sens réels de production d'une entité du monde physique ou sonore, comme aux sens figurés où l'activité humaine induite produit un excès; plus loin, ils peuvent être des instruments de transformation d'une source d'énergie ou des instruments dont l'énergie est apportée par le moulin pour réaliser une action; et enfin ils peuvent être des instruments dont la dénomination n'est pas explicite mais dont on sait qu'ils sont constitués par un élément donné (ces derniers relèvent plus en effet de la description). Il apparaît que le rôle télique et la qualité sémantique de l'argument sont déterminants et suffisants dans une majorité de réalisations `[transformer(entrée : produit naturel], créer(sortie : artefact, NParole(B2), Nparole(B2>)]`, et que le rôle constitutif soit nécessaire pour compléter le filtrage pour les autres réalisations.

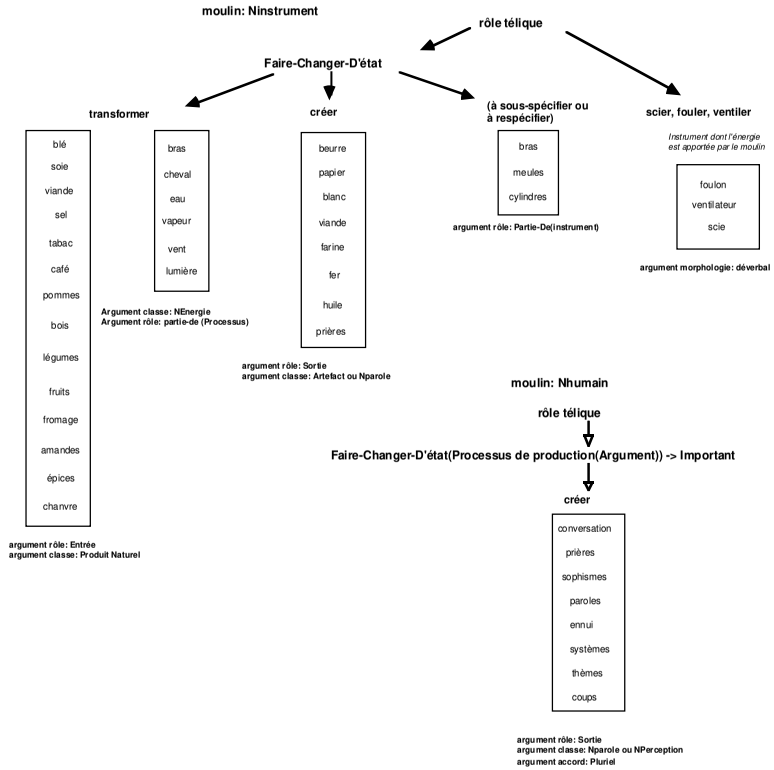


Figure 4.24 : `Moulin`, des préférences d'interprétation.

Peut-on y voir une préférence d'interprétation qui se développe en parallèle de la plus vers la moins conceptuellement définie et de la moins vers la plus descriptivement définie ? Le problème de la préférence d'interprétation peut d'ailleurs être envisagé au cours de l'analyse par une confrontation dynamique des savoirs représentés.

Si `moulin à bras` appartient à deux classes sémantiques dont l'une est plus productive en termes de réalisations attestées, il ne faut pas pour autant laisser de côté l'interprétation "la plus faiblement" réalisée. Si telle interprétation est plus saillante, il semble cohérent de la privilégier mais il est impératif de récupérer les autres interprétations possibles. On peut d'ailleurs considérer que la représentation sémantique attachée à `bras` s'accorde avec cette volonté de représenter le flexible de la langue qui se matérialise dans ce cas par une conversion de type dans une des interprétations possibles pour `moulin à bras`. On peut ainsi considérer que l'interprétation `A2 [rôle télique : transformer, argument rôle : Nénergie]` soit initiée par la représentation sémantique attachée à `moulin` puis résolue par la représentation sémantique attachée à `bras` qui permet une conversion de type. On construit ainsi les deux interprétations possibles : des informations supplémentaires devraient permettre de choisir l'une ou l'autre des interprétations. De la même manière, on peut associer à `cheval` un rôle constitutif qui exprime le fait qu'un cheval peut être un composant d'un processus de traction, de poussée, etc.; `cheval` devient dans ce cas par conversion un `Nénergie`.

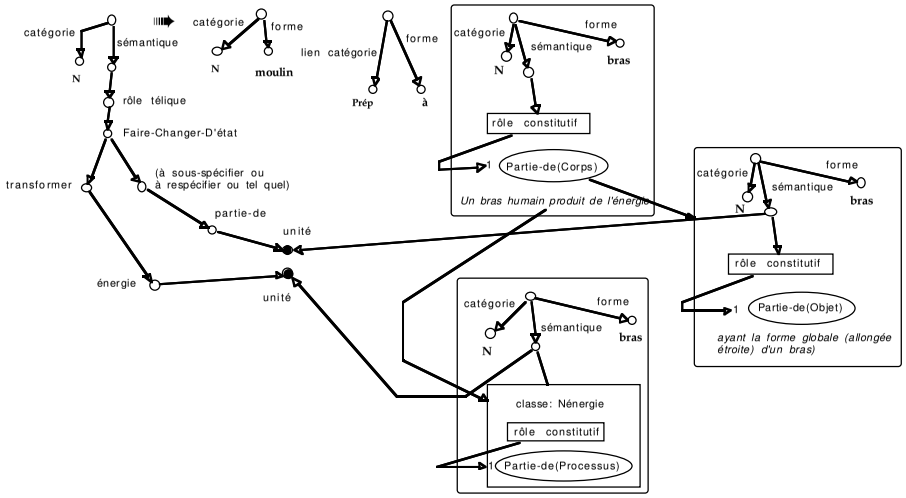


Figure 4.25 : `Moulin à bras` et préférences d'interprétation.

4.5.2.7. `Moulin` et les noms d'instruments

Etendre la représentation sémantique aux noms d'instruments : comment mettre en oeuvre une catégorisation plus générale tout en préservant une évolution potentielle de cette représentation initiale? Cette représentation sémantique tend à identifier (à isoler) les comportements sémantiques que l'on peut attacher à `moulin` et qui permettent de construire des (débuts d') interprétations possibles qu'il reste possible d'affiner. Ces comportements traduisent-ils des opérations sémantiques plus générales qu'il serait possible de factoriser pour une classe de noms particuliers (en particulier pour les noms d'instrument). La figure suivante illustre ce que pourrait être cette recatégorisation.

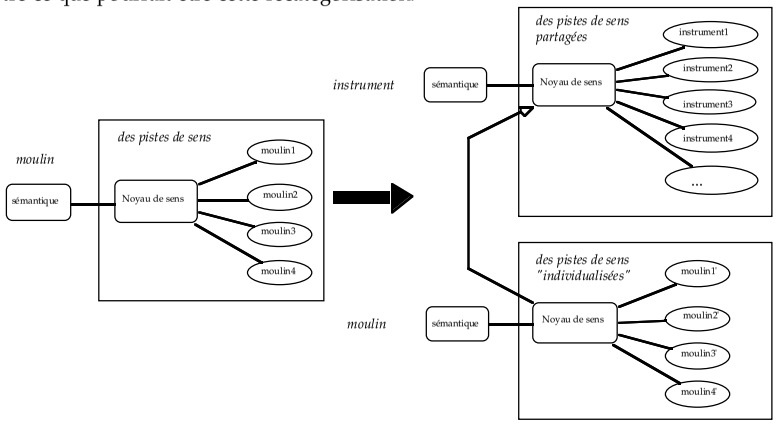


Figure 4.26 : `Moulin` et noms d'instrument, des comportements sémantiques partagés.

Si une approche prototypique se révèle adéquate pour rendre compte d'un micro-phénomène particulier, il convient de ne pas perdre de vue la place de ce micro-phénomène dans un

système plus complexe, ou tout au moins de dégager de l'étude particulière de celui-ci, les éléments qui le dépassent.

4.5.2.8. Construire un espace de concepts

Si les concepts n'existent qu'au travers de nos activités de langue, l'approche retenue s'accorde, semble-t-il, assez bien avec la construction d'un espace de concepts qui eux-aussi se définiraient par touches successives. L'approche montante poursuivie à partir de représentations prototypes sous-déterminées vise à construire une représentation sémantique des unités linguistiques manipulées par affinements successifs; l'espace conceptuel à venir doit donc se mettre en place à mesure que la construction de l'espace sémantique se développe. Ce que montre le travail de D. Kayser est qu'il est impossible d'anticiper toutes les significations que l'on peut attacher à un mot, et qu'au contraire celles-ci se distinguent au cours de la construction dynamique d'une interprétation. Il est cependant nécessaire de disposer de moyens efficaces de pouvoir ajuster les représentations qui sont faites dans une modélisation des activités d'interprétation. De la même manière que la figure précédente tend à organiser, pas à pas, le savoir sémantique (en organisant une généralisation minimale et un savoir particulier modifiable), il faut rendre possible la mise en place d'un espace conceptuel qui s'articule au mieux à ce qui est en cours de construction.

4.6. Comment nomme-t-on les choses ?

"Si tu vois quelque chose de loin et ne comprends pas de quoi il retourne, tu te contenteras de le définir comme un corps en extension. Quand il se sera approché de toi, tu le définiras alors comme un animal, même si tu ne sais pas encore s'il s'agit d'un cheval ou d'un âne. Et enfin, quand il sera plus près, tu pourras dire que c'est un cheval, même si tu ne sais pas encore si c'est Brunel ou Favel. Et seulement quand tu seras à la bonne distance, tu verras que c'est Brunel (autrement dit ce cheval et pas un autre, quelle que soit la façon dont tu décides de l'appeler)." (Umberto Eco, Le nom de la rose, Livre de Poche 1982).

Quand on ne sait pas quelle désignation précise donner aux choses auxquelles nous sommes confrontées, appréhende-t-on ces éléments dans une globalité indéfinie, quitte à préciser ultérieurement ce flou initial? Ou bien tente-t-on de repérer les spécificités de ces éléments qui nous permettront de l'identifier? Pour nommer, classifier, il faut voir les choses de près, identifier les éléments, composants de celles-ci, pour les comparer puis les apparier ou ranger dans les classes déjà connues. Si la spécificité d'une chose qui permet de la désigner précisément n'est pas directement perceptible, on est dans la situation d'un observateur éloigné de la chose à identifier et qui ne peut reconnaître celle-ci de manière certaine. De plus, suivant les savoir-faire, les points de vue, les domaines d'expérience, les choses à nommer ne seront pas appréhendées de manière identique d'un individu à un autre : les hommes (mâles), les kangourous, les opossums et les chauves-souris, l'arc-en-ciel sont désignés par une même expression par les arborigènes d'Australie, de même qu'une fleur, le maïs et le bison chez les Osages. L'étude de la catégorisation devrait contribuer à éclairer le questionnement constant sur ce *"qui permet à la pensée d'opérer sur les êtres une mise en ordre, un partage en classes, un groupement nominal par quoi sont désignées leur similitudes et leurs différences,..., là où depuis le fond des temps le langage s'entrecroise avec l'espace"* (Michel Foucault 1966).