

## Quels types de connaissance sémantique pour Questions-Réponses ?

Vincent Barbier  
LIMSI, UPR CNRS 3251, Bât. 508  
Université de Paris XI, 91403 Orsay  
barbier@limsi.fr

### Mots-clefs – Keywords

Système de Question Réponse, ressources sémantiques, évaluation des reformulations  
Question Answering system, semantic resources, evaluation of reformulations

### Résumé - Abstract

Les systèmes de Questions Réponse ont besoin de connaissances sémantiques pour trouver dans les documents des termes susceptibles d'être des reformulations des termes de la question. Cependant, l'utilisation de ressources sémantiques peut apporter un bruit important et altérer la précision du système. ne fournit qu'une partie des reformulations possibles.

Cet article présente un cadre d'évaluation pour les ressources sémantiques dans les systèmes de question-réponse. Il décrit la fabrication semi-automatique d'un corpus de questions et de réponses destiné à étudier les reformulations présentes entre termes de la question et termes de la réponse. Il étudie la fréquence et la fiabilité des reformulations extraites de l'ontologie WordNet.

Question Answering systems need semantic knowledge to find in the documents terms that are reformulations of the question's terms. However, the use of semantic resources brings an important noise and a system's precision might get depreciated.

This article presents a framework for evaluating semantic resources in question-answering systems. It describes the semi-automated construction of a corpus containing questions and answers. This corpus can be used to study the various reformulations between the terms of the questions and the terms of the answers. This article studies the frequency and reliability of reformulations given by the WordNet ontology.

# 1 Introduction

La problématique questions-réponses (QR) consiste à concevoir un programme qui reçoit de l'utilisateur une question formulée en langue naturelle et recherche dans une collection de documents un fragment de texte, typiquement un court passage, répondant à cette question. Les campagnes d'évaluation TREC, dont l'enjeu est de répondre à des questions factuelles par de courts passages d'articles de journaux ont encouragé la réalisation de nombreux systèmes de QR et donné une évaluation globale de ces systèmes.

La plupart des systèmes sont constitués de trois étapes principales (Grau, 2004). Tout d'abord, l'analyse de la question qui permet de découper la question en termes et de trouver le type attendu de la réponse. Ensuite, la recherche des documents dans la collection, au moyen d'un moteur de recherche. On effectue pour cela une ou plusieurs requêtes successives composées des termes de la question et éventuellement de reformulations de ceux-ci. Enfin, la sélection des réponses mettant en jeu des critères de pertinence faisant intervenir la syntaxe et la sémantique.

Pour trouver des documents pertinents, les systèmes ont besoin de reconnaître dans les documents des reformulations de termes de la question. Par exemple, dans la collection d'articles de journaux Aquaint (TREC11), la question "What is the name of the volcano that destroyed Pompeii ?", ne possède pas de réponse réutilisant le verbe *to destroy*. Par contre, des équivalents comme *to devastate*, *to bury under ashes...* permettent de retrouver la réponse.

Les reformulations peuvent intervenir à deux endroits. Premièrement, au niveau de la requête, pour permettre de retrouver les documents qui n'utilisent pas la même formulation de la question, c'est-à-dire, augmenter le rappel de l'étape de recherche des documents. Ensuite au niveau de la sélection des documents, des passages ou des phrases retournés, afin d'améliorer leur classement et sélectionner ceux dont la formulation approche le mieux possible celle de la question. Plus les reformulations sont complexes et plus elles sont susceptibles d'engendrer du bruit, c'est-à-dire d'amener le système à juger pertinents des documents ou passages qui ne le sont pas. Les systèmes de questions-réponses doivent donc utiliser les ressources linguistiques avec précaution, et pour cela, doivent estimer la fiabilité des reformulations utilisées.

Notre but est d'étudier le besoin de reformulations des systèmes de questions réponses, de chercher quels types de reformulations seraient utiles à ces systèmes et quel est le degré de fiabilité de ces reformulations. Cette étude se fonde sur un corpus de questions réponses constitué automatiquement. Après un état de l'art des systèmes de questions réponses et de la façon dont ils utilisent les ressources sémantique, nous décrivons la construction du corpus, puis nous étudierons les reformulations obtenues grâce à l'ontologie WordNet, et tenterons d'évaluer leur fiabilité.

## 2 La sémantique dans les travaux actuels

Lors du filtrage des documents ou des passages, les systèmes utilisent différents critères pour évaluer la présence d'une réponse. Ceux-ci se répartissent en critères syntagmatiques et paradigmatiques. Les critères paradigmatiques sont ceux qui déterminent une proximité entre un élément de la question et un élément de la réponse. Les critères syntagmatiques sont ceux qui tiennent compte de la disposition de ces éléments dans les entités textuelles où ils apparaissent, ou de leurs dispositions relatives. Les systèmes allient ces deux types de contraintes sous

## Quels types de connaissance sémantique pour Questions-Réponses ?

diverses formes.

Des techniques les plus restrictives aux plus souples, la contrainte syntagmatique peut être implémentée par :

- L'utilisation de patrons d'extraction de la réponse, par exemple une date de naissance peut être repérée par "X was born in <date-naissance>" ou "X ( <date-naissance> - <date-décès> )". (Jijkoun & de Rijke, 2004) montre que cette méthode apporte des résultats pour des questions simples (TREC8). Elle est de plus bien adaptée à des collections de documents de taille très importante, comme celle rendue disponible par la toile.
- Une preuve logique visant à démontrer une implication possible entre deux graphes, représentant respectivement la question et un passage candidat à contenir la réponse (Moldovan *et al.*, 2002). Cette technique consiste à écrire la question et un passage candidat sous forme logique. Par exemple "How did Hitler die?" est retranscrit par :

manner\_at(e) & hitler\_nn(x) & die\_vb(e,x,y).

La preuve peut utiliser des axiomes linguistiques et sémantiques. Un axiome sémantique, tiré de WordNet, pourrait être "to kill cause to die" ce qui a été retranscrit par :

kill(e1,x1,x2) -> die(e1,x2,\$c) ou encore - kill(e1,x1,x2) | die(e1,x2,\$c)

où \$c est un argument indéterminé, et '-' l'opérateur de négation.

- Une mesure de la proximité spatiale des éléments de la question dans le passage candidat.

En ce qui concerne la contrainte paradigmatique, le but est d'associer les éléments de la question à des éléments correspondants dans la réponse. Il peut s'agir de repérer dans le passage candidat un terme correspondant au type attendu de la réponse (entités nommées trouvées dans une taxonomie ou concept d'une ontologie), ou de repérer des reformulations sémantiques des termes de la question.

La mise en correspondance des termes de la question et de la réponse nécessite des ressources. Le type de ressource majoritairement utilisé sont les dictionnaires de synonymes (Magnini *et al.*, 2002; Ferret *et al.*, 2001). Mais il est possible d'utiliser les différentes relations d'une ontologie telle que WordNet (Moldovan *et al.*, 2002). L'ontologie WordNet (Fellbaum, 1998) est intéressante pour sa couverture générale de la langue. C'est sur elle que s'axera en premier lieu notre travail.

WordNet est un lexique dont les entrées, les synsets, sont assimilables à des concepts. Les catégories grammaticales traitées sont les noms, les verbes, les adjectifs et les adverbes. La plupart des travaux concernant WordNet se concentrent sur l'ontologie des noms. Les entrées de WordNet sont décrites par leurs connexions avec les autres entrées du lexique. Ces liens sont de plusieurs types : l'hyponymie, qui relie un concept général à des concepts plus spécifiques, et son inverse, l'hypéronymie. Cette paire de relation constitue la principale structuration de l'ontologie noms. Les noms sont aussi organisés selon la relation d'holonymie qui relie un élément à un ensemble le contenant (par exemple *argentin* est un méronyme de *Argentine*). Les relations structurant les verbes sont le couple hyponymie/hypéronymie, ainsi que les relations de cause (ex : tuer cause le fait de mourir) et d'implication (ex : ronfler entraîne que l'on dort). Les adjectifs, peu structurés, sont définis par leurs synonymes et antonymes, et surtout par les noms auxquels ils se rapportent.

A la structuration des liens, il faut ajouter les définitions (glosses), qui peuvent elles aussi apporter des liens utiles qui ne seraient pas codés explicitement.

LCC fait un usage important de ces liens grâce à une version étendue de WordNet où les définitions sont codées sous forme de graphes et les mots de ces définitions ont été désambiguïsés manuellement, c'est-à-dire rattachés au synset leur correspondant le mieux.

Cependant, les relations codées dans WordNet ne sont qu'une partie des relations existantes. Par exemple, les liens de WordNet donnent principalement des relations entre mots d'une même catégorie grammaticale. Or certaines reformulations ne s'effectuent pas mot à mot mais sur une combinaison de deux ou plusieurs mots (Jacquemin, 1999). Certaines constituent une transformation complète de la phrase. Ces reformulations complexes peuvent remplacer un mot par un mot d'une catégorie grammaticale différente. Par exemple la question :

“What is the democratic party *symbol*?”

possède une réponse de la forme :

“The democratic party was *represented* as a donkey.”

Pour remédier aux manques de WordNet, on peut utiliser les relations morphologiques qui relient par exemple le verbe *to make* au nom *maker*. Cependant ces relations morphologiques couvrent le lexique de façon lacunaire et hasardeuse. Ainsi, en anglais, les termes décrivant les relations conjugales : “husband, wife, married” ne sont pas morphologiquement reliés. Pour combler ces manques, (Claveau & Sebillot, 2004; de Chalendar G., 2001) acquièrent automatiquement des relations nom-verbe sur corpus.

Plusieurs sources de connaissances sémantiques sont donc disponibles, synonymes, ontologies, ou relations apprises sur corpus. L'architecture générale des systèmes de QR est définie et relativement homogène d'un système à l'autre. Dans ce cadre, il est maintenant nécessaire d'examiner plus en détail l'apport des différentes étapes, ainsi que d'évaluer l'apport des ressources linguistiques utilisées.

### 3 Construction du corpus

Afin d'étudier l'utilité et d'évaluer la fiabilité des reformulations, nous sommes passés par la construction d'un corpus. Nous voulons pouvoir utiliser ce corpus pour étudier, pour chaque terme d'une question, les termes qui lui correspondent dans les réponses. Ce corpus pourra aussi servir pour apprendre automatiquement des mesures de proximités sémantiques.

Actuellement, ce corpus est fondé sur 54 questions de la campagne d'évaluation TREC11. Nous l'étendrons à plus de questions. La collection de documents utilisée est la collection Aquaint fournie pour cette campagne.

A chaque question est associé un ensemble de documents pertinents et non pertinents. Un passage est jugé pertinent s'il contient la réponse à la question, laquelle est repérée par une expression régulière, telle que “John( F.)? Kerry”, qui indique que les caractères “ F.” sont facultatifs. Cette méthode laisse passer du bruit dans la liste de documents pertinents, mais elle permet d'automatiser la sélection de ces documents.

Quels types de connaissance sémantique pour Questions-Réponses ?

### 3.1 Eviter les biais

Pour la constitution du corpus, il est nécessaire d'effectuer une requête sur un ensemble de documents en évitant que cette requête n'introduise un biais par rapport au sujet de l'étude. Dans notre cas l'étude porte sur les reformulations des mots de la question, et notamment, les reformulations obtenues en suivant les liens proposés par WordNet. Son but est de mesurer la fiabilité de certaines reformulations par rapport à d'autres, ou plus précisément de certaines caractéristiques de ces reformulations.

Nous avons opté pour l'utilisation de requêtes booléennes plutôt que vectorielles car ces premières permettent de mieux contrôler le contenu des documents rapportés.

Or, si la requête se présente sous la forme d'une conjonction de termes, ces termes apparaîtront nécessairement dans les documents pertinents et les documents non pertinents. Ceci rendrait inutilisable une mesure fondée sur la comparaison des fréquences d'un terme entre documents pertinents et non pertinents.

On peut tenter de pallier ce problème en présentant les requêtes sous la forme d'une conjonction de disjonctions où chaque terme de la question est remplacé dans la requête par une disjonction de variantes de ce terme. Cependant, les variantes choisies sont favorisées par rapport aux autres. De plus, si ces variantes sont extraites d'une ressource telle que WordNet, les reformulations apparaissant dans les documents collectés seront principalement celles de cette ressource, ce qui réduirait la portée de l'étude à un type de relations très particulières.

### 3.2 Requête et filtrage

La méthode que nous avons retenue est la création à partir des termes de la question d'une "requête à trou", c'est-à-dire une requête où au moins un des termes ne sera pas représenté. Les termes omis seront l'objet de l'étude. Les autres termes sont reformulés grâce WordNet. Les variantes autorisées sont les synonymes des termes de la question, ainsi que les mots le plus courant des synsets situés à un lien de distance. L'étendue des reformulations est volontairement réduite afin que la requête engendrée ne rapporte pas trop de documents non pertinents. Une plus grande expansion de la requête augmenterait le bruit qui est déjà important (CF Table 2). Enfin, les termes les moins significatifs de la question ne sont pas utilisés dans les requêtes. Cette significativité est estimée par un expert humain qui classe les termes de la question par ordre décroissant. Par exemple, les termes de la question "What is the name of the volcano that destroyed the ancient city of Pompeii?" ont été classés dans l'ordre suivant :

Pompeii > volcano > destroy > ancient > city.

De manière systématique, les entités nommées sont considérés comme de meilleurs filtres que les noms communs et les autres catégories grammaticales décrites dans WordNet.

Cet ordre permet ensuite de créer des requêtes adéquates de façon automatique. Par exemple, si l'on veut étudier les reformulations possibles du terme "destroy", on utilisera pour la requête :

Pompeii & expansion(volcano) & expansion( ancient )  
avec expansion( volcano) = (mountain | mount | crater | volcano)

Une fois les documents rapportés, on sépare les documents pertinents des non pertinents. Il faut noter que par cette méthode, les documents pertinents sont obtenus par une même requête, ce qui réduit les risques de biais.

Certains patrons de réponses sont des filtres très imprécis, par exemple une date ou un personnage souvent cité dans le corpus pour des raisons variées. Dans certains cas, la question contient des termes assez précis - par exemple, une entité nommée - pour permettre un filtrage efficace des documents. Mais si ce n'est pas le cas, il peut être nécessaire de rendre la requête plus précise en ajoutant des termes ou en réduisant l'étendue des reformulations. Dans le cas extrême, on pourra supprimer une question de l'étude.

### 3.3 Résultat

Le corpus obtenu contient 5952 documents jugés pertinents par le programme et 65419 documents jugés non-pertinents. Il a été testé manuellement pour 16 questions contenant 114 termes. Nous avons obtenu 295 documents pertinents, et 156 erreurs, ce qui correspond à un bruit de 35%. Un nettoyage manuel du corpus semble donc nécessaire.

patron +	patron -
5952	65419

Table1: taille du corpus brut.

	patron +	patron -
validé +	294(65%)	-
validé -	156(35%)	-

Table2: qualité du corpus validé.

Dans le sous-corpus validé, le nombre moyen de documents pertinents rapportés par question est de 10,6. Ce nombre est très variable d'une question à l'autre. Les extrêmes valent 1 et 43. Dans les deux quartiles centraux, ce nombre évolue de 3 à 11,5. Cette variabilité s'explique par le fait que certaines informations sont répétées (vie privée de stars, événements sportifs marquants) alors que d'autres sont très peu cités. D'autre part, on doit noter que les filtrages effectués pour réduire le bruit du corpus font nécessairement disparaître certains documents pertinents.

## 4 Quelle sémantique pour les systèmes de questions réponses

### 4.1 Les types de reformulations présentes.

A présent, analysons les types de reformulations présentes dans le corpus. Ceci peut s'effectuer manuellement en recensant et examinant les liens, ou de façon plus automatique, en mesurant pour différentes caractéristiques des liens, les fréquences d'apparition dans les phrases contenant la réponse, comparées aux fréquences d'apparition dans le reste du corpus.

#### 4.1.1 Etude manuelle des liens

Nous avons examiné les liens trouvés par le programme, et nous avons noté ceux qui semblaient pertinents. Pour qu'un lien nous semble pertinent il faut qu'un trait sémantique se transmette d'un bout à l'autre de la chaîne. On peut avoir plusieurs niveaux de pertinence, selon que les traits sémantiques partagés sont plus ou moins spécifiques. Par exemple, le lien

*volcano-(hyponyme)->Krakatau-(holonyme)->Indonésie-(hypéronyme)->country-(hyponyme)-> Philipines*

## Quels types de connaissance sémantique pour Questions-Réponses ?

possède une certaine pertinence, car il relie des noms géographiques, mais celle-ci est très faible.

Dans le corpus de test, on distingue deux catégories de liens intéressants : des liens fortement pertinents, pointant vers des mots susceptibles de faire partie de la réponse. Et d'autres plus distants, pointant vers des mots qui forment le thème général de l'article. Dans un article sur l'éruption du Vésuve on trouve des mots comme *earth*, *Naples*, *Campania*, qui viennent renforcer les champs thématiques de la question.

Cette observation, à vérifier sur un corpus plus important, nous indique que des relations assez distantes pourront peut-être aider au filtrage de la question.

**Différentes catégories de lien** Voici une courte description des types de liens trouvés dans le corpus. Aussi la majorité des liens trouvés concernent les noms. Les relations entre verbes sont moins fréquentes. Les relations liées aux adjectifs sont rarement pertinentes.

Nous notons une part importante de liens triviaux. Il s'agit de la reprise du mot de la question tel quel ou sous une forme fléchi. Les reformulations par synonyme sont également présentes. Pour certains mots, les reformulations morphologiques, permettant de changer de catégorie grammaticale sont importantes : par exemple pour la question

“When did the *shooting* in Columbine happen?”

Les reformulations les plus trouvées sont les verbes *to kill* et *to shoot* (260 occurrences dans le corpus) devant le nom *shooting* (50 occurrences).

Les relations d'hyponymie-hypéronymie et de méronymie-holonymie fournissent toutes deux des liens pertinents. Cette prépondérance des relations entre noms s'explique par le fait que les noms sont la catégorie grammaticale la plus présente dans la question, et que le réseau des noms contient plus de liens que celui des verbes.

Nous remarquons que les reformulations se font parfois par une succession d'une seule relation, par exemple, une succession d'hypéronymes :

*political\_party* -(hypéronyme)-> *organization* -(hypéronyme)-> *social\_group* -(hypéronyme)-> *group*

Mais des reformulations de ce type sont rares par rapport à celles reliant deux concepts frères, c'est-à-dire deux concepts pointant vers un même concept par une même relation. La relation empruntée est le plus souvent celle d'hypéronymie-hyponymie, mais on peut aussi trouver les relations d'holonymie-méronymie et d'implication comme dans les exemples ci-dessous.

Relation de co-méronymie correspondant à une proximité géographique :

*Vesuvius* -(holonyme)-> *Italy* <-(holonyme)- *Pompeii*

Verbes reliés par une relation de co-implication :

*destroy* -(hypéronyme)-> *smother* -(implique)-> *cover* <-(implique)- *entomb*

De nombreuses relations de co-hyponymie

*score* -(hypéronyme)-> *achieve* -(hypéronyme)-> *succeed* <-(hypéronyme)- *hit*  
*city* -(hypéronyme)-> *municipality* <-(hypéronyme)- *town*

Ces observations sont encourageantes et une analyse détaillée sur un corpus plus fourni permettra d'y puiser les éléments d'une heuristique.

## 4.2 Fiabilité des reformulations.

### 4.2.1 Protocole d'évaluation d'une heuristique

**Caractéristiques et heuristiques** Les liens de WordNet permettent de constituer des chemins plus ou moins longs pour relier des mots sensés posséder un lien sémantique. Cependant, plus on s'éloigne et plus on risque d'aboutir à un mot non pertinent pour la question posée. Afin de réduire le bruit engendré, on peut utiliser des heuristiques fondées sur les caractéristiques de la géométrie des chemins parcourus (Budanitsky, 1999).

Le corpus peut servir à évaluer la fiabilité d'heuristiques complexes ou de caractéristiques élémentaires des chemins, telle que leur longueur ou la présence d'une relation particulière. Ces mesures élémentaires peuvent permettre de justifier ou invalider des heuristiques plus complexes.

Nous décrivons ci-après un cadre commun pour l'évaluation d'heuristiques complexes et de caractéristiques unitaires.

**Evaluer la fiabilité des caractéristiques** On peut représenter heuristiques et caractéristiques sous la forme d'une fonction qui à chaque élément d'un ensemble (ici, l'ensemble des chemins) associe une valeur, soit booléenne, soit réelle.

Nous avons pour l'instant considéré des caractéristiques booléennes telles que "*est de longueur N*" ou "*Contient la relation R*". On calcule parmi les chemins vérifiant cette caractéristique, combien sont "pertinents", c'est à dire pointent vers un document pertinent. On calcule la précision de cette caractéristique dans le corpus qui vaut :

$$\frac{\text{Nb de liens pertinents vérifiant la caractéristique}}{\text{Nb total de liens vérifiant la caractéristique}}$$

La valeur de la précision calculée dépend du ratio entre nombre de documents pertinents et non-pertinents. Elle n'a donc pas de réalité absolue mais permet d'effectuer des comparaisons entre deux caractéristiques ou heuristiques obtenues par le même protocole.

**Estimer la significativité des mesures** Nous appliquerons à notre protocole une méthode de validation par rééchantillonnage utilisée par exemple dans (Voorhees, 2002) pour la campagne d'évaluation TREC11. Cette méthode permettra d'estimer la variance des mesures effectuées et de savoir si elles sont fiables. Pour l'instant nous n'avons pas effectué cette étape de validation.

### 4.2.2 Les caractéristiques testées

Nous avons commencé par tester des caractéristiques élémentaires des chemins.

**Longueur** La première est la longueur. Le résultat est prévisible. En effet, on s'attend à ce que plus un chemin est de longueur courte et plus il est fiable que le protocole d'évaluation.

longueur	0	1	2	3	4
précision(%)	17*	40	20	10	5

Table3: fiabilité des reformulations en fonction de la distance dans WordNet.

\* Présence d'un biais dans le corpus pour les reformulations de longueur nulle.



## Quels types de connaissance sémantique pour Questions-Réponses ?

Ce résultat est très intéressant car il montre bien que la fiabilité d'un lien décroît avec sa longueur.

**Effets de la combinaison de plusieurs relations** Il semble que les relations apparaissent préférentiellement accompagnées de leurs relation opposée, par exemple, la relation d'hypéronymie accompagnée d'une ou plusieurs relation d'hyponymie. Nous voulons vérifier si certaines relations "vont bien ensemble" en termes de fiabilité de la reformulation obtenue, ou si d'autres assortiments ont des effets négatifs. Ce problème revient à mesurer par le protocole que nous avons présenté la fiabilité des deux caractéristiques booléennes suivantes :

**Cooccurrence exclusive** : vraie s'il y a présence dans un chemin de deux types de relation à l'exclusion de tout autre type

**Cooccurrence non exclusive** : vraie en cas de présence dans un chemin de ces deux types de relation qu'il y ait ou non d'autres types de relations présents.

Nous nous sommes limité aux cooccurrences non exclusives de 2 relations :

	hype	hypo		mero	holo
hype	6,4%	7,8%	mero	3.6%	6,3%
hypo	7,8%	4,9%	holo	6,3%	*40%

\*non significatif

Table4: fiabilité des reformulations selon les liens qui les composent.

La cooccurrence d'un type de relation avec lui-même n'est vérifiée que si cette relation est présente deux fois dans le chemin.

En ce qui concerne les cooccurrences d'hyponymies et d'hypéronymies, on constate que les chemins contenant uniquement l'une ou l'autre relation sont moins fiables que ceux contenant les deux. Une interprétation possible de ce résultat est que les chemins qui ne sont constitués que de l'une de ces deux relations causent un changement du niveau de généralité entre un terme et sa reformulation. Au contraire une reformulation mixte a plutôt tendance à conserver le niveau de généralité.

Ces résultats rejoignent les observations sur le corpus qui soulignaient la présence de reformulations mettant en relation des concepts "frères".

### 4.3 Utilité d'une catégorie de reformulation

Notre travail présente une méthode pour évaluer la fiabilité de différentes catégories de reformulations. Le problème suivant est d'évaluer l'utilité de ces catégories de reformulations. Il faut pour cette étude définir ce qu'est l'utilité d'une catégorie.

Une première piste est de considérer qu'une variante d'un terme est utile si le fait de la prendre en compte permet de trouver au moins un document pertinent alors qu'ignorer cette variante n'aurait pas permis d'en trouver. Le recensement de ces variantes utiles permettra d'évaluer dans quelle mesure des catégories de variantes sont utiles à la recherche des documents pertinents.

## Conclusion

Cette étude montre la possibilité d'utiliser un corpus de questions réponses constitué semi-automatiquement pour étudier les reformulations dans le cadre des systèmes de questions réponses.

Les résultats obtenus peuvent nous guider dans le choix de reformulations fiables. Ils semblent notamment indiquer l'utilité et la fiabilité de reformulations par mots "frères". Un autre facteur important pour juger de l'utilité d'un type de reformulations est la fréquence avec laquelle celle-ci peut s'utiliser. Ces considérations nécessitent une étude de fréquence sur corpus que nous effectuerons dans nos travaux futurs.

Notre travail a porté sur WordNet car c'est une ressource de large couverture lexicale et fortement structurée. Cependant, le corpus produit n'est pas biaisé en faveur de WordNet, et permettra par conséquent l'étude d'autres ressources sémantiques. Une perspective de mon travail est d'étudier l'utilité de ressources sémantiques constituées de liens nom-verbe et acquises automatiquement sur corpus par (de Chalendar G., 2001).

## Références

- BUDANITSKY A. (1999). *Lexical semantic relatedness and its application in natural language processing*. Rapport interne CSRG-390, University of Toronto, Department of Computer Science.
- CLAVEAU V. & SEBILLOT P. (2004). Extension de requêtes par lien sémantique nom-verbe acquis sur corpus. In *Actes de TALN 2004*.
- DE CHALENDAR G. (2001). *SVETLAN', un système de structuration du lexique guidé par la détermination automatique du contexte thématique*. PhD thesis, Université Paris XI.
- FELLBAUM C. (1998). *WordNet, an electronic lexical database*. Cambridge, MA: The MIT Press.
- FERRET O., GRAU B., HURAUPT-PLANTET M., ILLOUZ G. & JACQUEMIN C. (2001). Utilisation des entités nommées et des variantes terminologiques dans un système de question-réponse. In *Actes de TALN 2001*.
- GRAU B. (2004). Les systèmes de question réponse. In M. IHADJADENE, Ed., *Méthodes avancées pour les systèmes de recherche d'informations*. Paris: Hermes Sciences.
- JACQUEMIN C. (1999). Syntagmatic and paradigmatic representation of term variation. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL'99)*, p. 341–348.
- JIJKOUN V. & DE RIJKE M. (2004). Information extraction for question answering : Improving recall through syntactic patterns. In *Proceedings of ACL 2004*.
- MAGNINI B., NEGRI M., PREVETE R. & TANEV H. (2002). Mining knowledge from repeated co-occurrences: Diogene at trec 2002. In *Proceedings of the 11th Text REtrieval Conference, (TREC11)*.
- MOLDOVAN D., HARABAGIU S., GIRJU R., MORARESCU P., LACATUSU F., NOVISCHI A., BADULESCU A. & BOLOHAN O. (2002). Lcc tools for question answering. In *Proceedings of the 11th Text REtrieval Conference (TREC 2002)*.
- VOORHEES E. (2002). Overall view of the question answering track. in special publication. In *The 11th Text REtrieval Conference (TREC 2002)*.