

---

# Utilisation de la syntaxe pour valider les réponses à des questions par plusieurs documents.

Véronique Moriceau<sup>\*,\*\*</sup> — Xavier Tannier<sup>\*,\*\*</sup> — Brigitte Grau<sup>\*,\*\*\*</sup>

<sup>\*</sup> *LIMSI-CNRS, Orsay, France*

<sup>\*\*</sup> *Université Paris-Sud 11, Orsay, France*

<sup>\*\*\*</sup> *ENSIIE, Evry, France*

---

*RÉSUMÉ. Cet article présente FIDJI, un système de questions-réponses pour le français, combinant des informations syntaxiques sur la question et les documents avec des techniques plus traditionnelles du domaine, telles que la reconnaissance des entités nommées et la pondération des termes. Notamment, nous expérimentons dans ce système la validation des réponses dans plusieurs documents, ainsi que des techniques spécifiques permettant de répondre à différents types de questions (comme les questions attendant des réponses multiples (liste) ou une réponse booléenne).*

*ABSTRACT. This article presents FIDJI, a question-answering system for French, combining syntactic information with traditional QA techniques such as named entity recognition and term weighting. Among other uses of syntax, we experiment in this system the validation of answers through different documents, as well as specific techniques for answering different types of questions (e.g. yes/no or list questions).*

*MOTS-CLÉS : Systèmes de questions-réponses, analyse syntaxique, validation des réponses.*

*KEYWORDS: Question-answering, syntactic analysis, answer validation.*

---

## 1. Introduction

Cet article présente le système FIDJI<sup>1</sup> (Finding In Documents Justifications and Inferences), un système de questions-réponses en domaine ouvert pour le français.

Le projet dans son ensemble a pour objectif de valider des réponses en vérifiant que toutes les informations données dans une question sont bien retrouvées dans les passages de texte supportant la réponse. Cette vérification repose sur le fait de retrouver les différentes entités précisées dans la question correctement reliées entre elles, soit dans une même phrase, un même passage, ou dans plusieurs documents.

Lorsqu'une information est recherchée, elle peut être présente sous différentes formulations et peut requérir, pour être reconnue, l'utilisation de bases de connaissances sémantiques et la réalisation d'inférences avec plusieurs pas de raisonnement pour relier les différentes parties de la réponse. Or, si l'on dispose bien de bases lexicales contenant des variations sur les termes (synonymes par exemple), des bases conceptuelles permettant de relier les concepts pour décrire des événements ne sont pas disponibles pour le français et on ne peut donc envisager une analyse sémantique.

Aussi, notre propos consiste à retrouver les informations précisées par la question en se reposant sur la syntaxe, et notamment la reconnaissance des dépendances entre syntagmes et leurs différentes paraphrases possibles pour identifier les relations cherchées. Nous décrivons dans cet article le fonctionnement général du système, de l'analyse des questions à l'extraction de la réponse. Nous détaillons et évaluons notamment une technique permettant de valider une réponse à travers plusieurs documents.

L'article présente tout d'abord la stratégie générale mise en œuvre dans le système FIDJI (section 2), puis l'analyse syntaxique (section 3) et l'extraction et la justification de la réponse (section 4). Enfin, la section 5 décrit une évaluation et en étudie les résultats.

## 2. Présentation de FIDJI

La plupart des systèmes de questions-réponses peut extraire une réponse à une question factuelle quand celle-ci est explicitement présente dans le texte, mais ils ne sont pas capables de combiner plusieurs informations pour produire une réponse. FIDJI (Finding In Documents Justifications and Inferences), un système de questions-réponses en domaine ouvert pour le français, vise à introduire des mécanismes de compréhension reposant sur des inférences.

L'objectif est de produire des réponses qui sont entièrement validées par des extraits de textes (des passages).

---

1. Ce travail est financé partiellement par le projet CONIQUE ANR-05-BLAN-008501 et par OSEO dans le cadre du programme Quaero.

Valider les réponses à des questions.

La principale difficulté est qu'une réponse (ou des informations composant une réponse) peut être validée par plusieurs documents. Par exemple :

*Question : Quel premier ministre français s'est suicidé ?*

*Réponse : Pierre Bérégovoy*

*Passage 1 : Le premier ministre français Pierre Bérégovoy a mis en garde Bill Clinton contre...*

*Passage 2 : Deux ans plus, Pierre Bérégovoy s'est suicidé après avoir été impliqué...*

Dans cet exemple, la réponse *Pierre Bérégovoy* ne peut être entièrement validée par un seul passage : les informations *premier ministre français* et *s'est suicidé* sont validées par deux passages différents. En fait, la réponse est ici l'intersection de deux ensembles de réponses obtenues aux deux questions *Qui s'est suicidé ?* et *Qui sont les premiers ministres français ?* Dans ce cas, pour pouvoir proposer une réponse entièrement validée, il est nécessaire de décomposer la question en deux sous-questions.

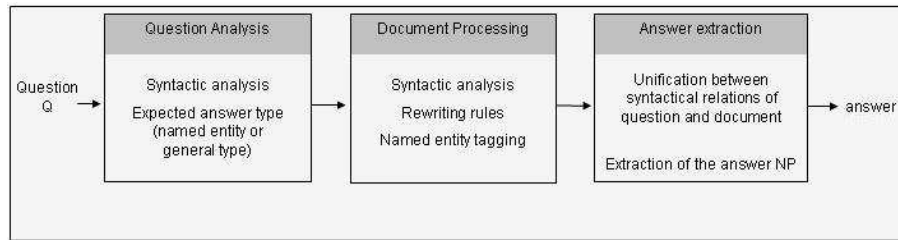
Une analyse syntaxique peut fournir de telles décompositions pour les questions. Beaucoup de systèmes de questions-réponses utilisent des informations syntaxiques, en particulier les relations de dépendance, principalement pour l'extraction des réponses. Deux approches émergent : la première consiste à rechercher un appariement exact entre les relations de dépendance de la question et celles d'un passage [KAT 03], tandis que la seconde approche calcule une distance d'édition entre les arbres représentant la question et le passage [LIG 07].

[KAT 05] propose une stratégie pour décomposer les questions à un niveau syntaxique et sémantique : ceci permet à leur système de rechercher des informations dans plusieurs ressources. Il utilise un certain nombre d'"annotations paramétrées" et de patrons sémantiques appliqués à toute la collection de documents afin de relier une question aux informations d'un ou plusieurs documents. Ce système est principalement construit pour répondre aux questions portant sur des objets ou des propriétés (par exemple, une date de naissance, la population d'une ville, etc.). Le système de questions-réponses IRSAW pour l'allemand [HAR 08] adopte lui aussi une stratégie de décomposition des questions en s'appuyant sur un analyseur syntactico-sémantique.

Notre but est d'extraire et de valider des réponses en allant au-delà d'un appariement syntaxique exact entre la question et le passage, et cela sans utiliser de ressources sémantiques. Dans ce contexte de validation de réponse, nous avons remarqué que la stratégie de validation à appliquer (validation par un seul ou plusieurs documents) peut être guidée par la question, et en particulier par le type de réponse attendu. En effet, beaucoup de questions factuelles attendent une réponse d'un type qui peut être :

– une entité nommée comme dans *Qui a obtenu le Prix Nobel de la paix en 1995 ?* qui attend une réponse de type PERSONNE ;

– un type plus précis comme dans *Quel président russe a assisté au G7 en 2007 ?*, qui attend aussi une réponse de type PERSONNE mais dont le type est précisé explicitement dans la question (*président russe*). Le type précis n'est pas issu d'une liste



**Figure 1.** Architecture de FIDJI

prédéfinie : il est identifié automatiquement lors de l'analyse de la question.

A la vue des différents exemples étudiés, nous faisons l'hypothèse que le type de la réponse est un élément qui peut être validé dans des documents autres que ceux d'où la réponse est effectivement extraite. Dans cet article, nous présentons cette stratégie. Ainsi, FIDJI utilise des informations syntaxiques, en particulier des relations de dépendances, qui vont permettre notamment de décomposer les questions. Le but est alors de vérifier la présence des relations syntaxiques de la question dans un passage justificatif et de confirmer le type de la réponse potentielle dans ce même passage ou dans un autre document, ceci afin de valider entièrement la réponse. La figure 1 présente l'architecture de FIDJI.

### 3. Analyse des questions et des documents

Notre système s'appuie sur les analyses syntaxiques produites par l'outil Syntex [BOU 00], un analyseur robuste pour le français. Syntex est utilisé pour analyser à la fois les questions et la collection de documents d'où sont extraites les réponses.

#### 3.1. Analyse des questions

L'analyse des questions consiste à identifier :

- les dépendances syntaxiques : elles sont fournies par Syntex ;
- le type de la question : factuelle, définition, booléenne, liste ;
- le type de réponse attendu : soit une entité nommée et/ou un type plus précis.

Les questions attendant des réponses de type liste sont celles qui contiennent explicitement un type de réponse au pluriel (par exemple *Quelles planètes... ?*, *Qui sont les... ?*, etc.).

Lors de l'analyse syntaxique de la question, la réponse à extraire est représentée par une variable (notée REPONSE) introduite dans les relations de dépendance. Le type de la question est principalement déterminé grâce à la forme de la question (pronom interrogatif, etc.). Le type spécifié de réponse attendu, s'il existe, est quant à lui le

Valider les réponses à des questions.

groupe nominal lié à la variable réponse par une relation *attribut* (équivalent à une relation *est\_un*). Par exemple :

Q : Quand fut construite la Tour Eiffel ?

Dépendances : DATE(REPONSE)

SUJ(construire, Tour Eiffel)

AUX(construire, être)

Type de la question : factuelle

Type de réponse attendu : DATE (entité nommée)

Q : Quelle déclaration fut adoptée par l'ONU en 1948 ?

Dépendances : attribut(REPONSE, déclaration)

AUX(être, adopter)

SUJ(adopter, REPONSE)

modif\_par(adopter, ONU)

DATE(1948)

Type de la question : factuelle

Type de réponse attendu : déclaration (type plus précis)

Dans le premier exemple, le type de la réponse est une entité nommée alors que dans le second exemple, c'est un mot ou groupe de mots directement extrait de l'analyse syntaxique de la question.

### 3.2. Analyse des documents

Notre approche consiste à déterminer, pour une question donnée, si toutes les caractéristiques de la question (en l'occurrence les dépendances syntaxiques) peuvent être trouvées dans un ou plusieurs documents. Dans ce but, FIDJI détecte les implications syntaxiques entre la question et les passages. Toute la collection de documents est donc également analysée syntaxiquement par Syntex.

Comme les informations dans les documents ne sont pas toujours exprimées de la même façon que dans les questions (par exemple, par le biais de variations syntaxiques), il est indispensable de raisonner sur les relations de dépendance syntaxique. De la même façon que [BOU 07], nous avons implémenté environ 40 règles de réécriture pour tenir compte, entre autres, des variations comme les changements voix passive/active, les nominalisations de verbes [JAC 96], les appositions, coordinations, etc. Ces règles de réécriture sont appliquées à toute la collection analysée.

De cette façon, quelle que soit la forme syntaxique de la question, le système est capable de trouver une formulation équivalente dans un passage justificatif : il est ainsi possible d'avoir un appariement exact soit entre les dépendances syntaxiques de

Véronique Moriceau, Xavier Tannier, Brigitte Grau.

$$\begin{array}{l} \text{SUJ(Verbe, NP1)} \\ \text{OBJ(Verbe, NP2)} \end{array} \Rightarrow \begin{array}{l} \text{SUJ(Verbe, NP2)} \\ \text{AUX(être, Verbe)} \\ \text{modif\_par(Verbe, NP1)} \end{array}$$

**Figure 2.** Exemple de règle de réécriture : voix active vers voix passive.

la question et d'un passage, soit entre les dépendances syntaxiques de la question et celles d'un passage obtenues par réécriture.

L'exemple suivant illustre la règle de réécriture pour la reformulation passif/actif (les relations obtenues par réécritures sont en italique).

Q : Quelle ville a été secouée par un tremblement de terre le 17 janvier ?

DATE(17 janvier)  
attribut(REPONSE, ville)  
SUJ(secouer, REPONSE)  
AUX(être, secouer)  
modif\_par(secouer, tremblement)  
attribut\_de(tremblement, terre)

Text : Le tremblement de terre qui a secoué, lundi 17 janvier, Los Angeles ne serait pas associé directement à la fameuse faille de San-Andreas...

SUJ(secouer, tremblement)  
OBJ(secouer, Los Angeles)  
*SUJ(secouer, Los Angeles)*  
*AUX(être, secouer)*  
*modif\_par(secouer, tremblement)*  
attribut\_de(tremblement, terre)  
DATE(17 janvier) ...

Les relations de dépendance réécrites sont obtenues par application de la règle de la figure 2. Dans cet exemple, toutes les relations de la question sont présentes dans l'analyse du passage.

### 3.3. Autres ressources : les entités nommées

Les entités nommées des documents sont étiquetées en utilisant un ensemble d'environ 160 types [ROS 07] (e.g. personne, organisation, lieu, nationalité, date, nombre, etc.). Cet étiquetage, associé aux résultats de l'analyse de la question, est utile pour vérifier que le type d'entité nommée attendu par la question est bien le même que celui de la réponse extraite dans le document. Par exemple, la question suivante attend une

Valider les réponses à des questions.

réponse de type LIEU :

Question : Où Barbara Hendricks a-t-elle donné son premier concert de l'année ?

```
LIEU(REPONSE)
SUJ(donner, Barbara Hendricks)
OBJ(donner, concert)
attribut_de(concert, année)
```

Passage étiqueté : <pers>Barbara Hendricks</pers> a donné son premier concert de l'Année nouvelle à <ville>Sarajevo</ville>.

```
SUJ(donner, Barbara Hendricks)
OBJ(donner, concert)
attribut_de(concert, année)
...
```

Dans cet exemple, toutes les relations de dépendance syntaxique de la question sont présentes dans le passage et le type de réponse attendu (LIEU) correspond bien au type de la réponse extraite *Sarajevo* (le type VILLE est un sous-type de LIEU dans la hiérarchie des entités nommées).

## 4. Recherche de la réponse

### 4.1. Sélection des passages

Les documents sont indexés par le moteur de recherche Lucene [HAT 04]. Les mots-clés utilisés pour interroger l'index sont l'ensemble des mots significatifs de la question<sup>2</sup> ; les 100 premiers documents sont conservés (ainsi, seuls ces documents sont considérés par la suite).

### 4.2. Extraction de la réponse

La stratégie employée pour l'extraction de la réponse dépend du type de la question. L'utilisation de l'analyse syntaxique se situe au niveau de la phrase. Les informations syntaxiques, lorsqu'elles sont utiles, sont combinées avec d'autres paramètres.

#### 4.2.1. Questions factuelles

Les phrases candidates sélectionnées sont celles qui comptent le plus de relations en commun avec la question. Une fois ces phrases obtenues, deux cas sont à considérer :

---

2. C'est-à-dire les mots étiquetés comme nom, verbe, adjectif ou adverbe par l'analyseur syntaxique.

1) Au moins une dépendance de la question comportant la variable 'REPONSE' s'unifie avec une dépendance de la phrase. Dans ce cas, cette variable est instanciée par un lemme, qui est considéré comme la tête de la réponse. La réponse complète est ensuite composée en ajoutant les modificateurs présents dans la phrase (compléments du nom et adjectifs). Si cette réponse est du type d'entité nommée attendu, un poids plus important lui est attribué (voir la section 4.4.1).

2) Si la variable 'REPONSE' ne trouve pas d'instanciation dans la phrase, des éléments ayant le type d'entité nommée approprié sont recherchés. Ceci a pour but de contrebalancer les erreurs d'analyse. Il arrive bien entendu régulièrement que la réponse soit présente dans le passage mais que les seules relations syntaxiques n'y conduisent pas.

Dans le cas d'absence d'entité nommée du type recherché dans la phrase considérée, elles peuvent être collectées dans les 2 phrases précédentes (avec un score moindre). Cette heuristique compense (très imparfaitement) l'absence d'une résolution anaphorique dans le système ; ceci produit inévitablement un certain bruit, mais qui est généralement masqué par la redondance des bonnes réponses (la fréquence d'extraction d'une réponse augmente en effet son score – section 4.4.1).

#### 4.2.2. *Questions listes*

L'analyse de Syntex, enrichie par certaines des règles de réécriture, fournit des informations sur la coordination des syntagmes.

Si une réponse est trouvée (par les techniques indiquées à la section précédente), les éventuelles relations de coordination avec cette réponse sont recherchées, et une liste est construite à partir des éléments ainsi collectés.

Si les réponses multiples sont préférées, les réponses atomiques sont retournées malgré tout ; en effet, une question étiquetée "liste" peut être satisfaite par un autre moyen. Par exemple, à la question *Qui sont les Dalton ?*, le pluriel indique certes que l'on attend une liste de noms, mais la réponse "*4 bandits*" est également tout à fait valable.

#### 4.2.3. *Questions booléennes*

Répondre à des questions booléennes est relativement proche de la tâche de validation de réponse, comme pratiquée par exemple lors des campagnes AVE (Answer Validation Exercise [ROD 08]). Dans cette tâche, les participants doivent considérer des questions, des réponses de systèmes et un texte de justification, et décider si la réponse à la question est à la fois correcte et validée entièrement par le texte.

Nous avons participé à AVE en langue française en 2008 : FIDJI s'est classé premier pour le français et second sur l'ensemble des candidats toutes langues confondues [MOR 08].



Valider les réponses à des questions.

Nous utilisons la même technique pour les questions booléennes : si la part de dépendances de la question trouvées dans la phrase dépasse un certain seuil (déterminé empiriquement à 70 %), la réponse retournée est 'oui'.

### 4.3. Validation du type de la réponse

Lorsqu'une phrase est sélectionnée et une réponse extraite, il arrive souvent que toutes les relations recherchées ne soient pas satisfaites. Dans certains cas, il est alors possible de vérifier la présence des dépendances manquantes dans d'autres documents. Dans l'état actuel du système, la seule validation effectuée par ce biais est celle du type précis de la réponse.

Ce type est fourni, lorsqu'il existe, par l'analyse de la question (voir section 3.1). Si le passage justificatif ne contient pas les informations permettant de le vérifier, une nouvelle question est construite pour valider la réponse candidate.

Dans notre exemple "*Quel premier ministre s'est suicidé...*", le type de réponse est *ministre* (tandis que le type d'entité nommée est PERSONNE). Le type *étendu* est *premier ministre*.

Q : Quel premier ministre s'est suicidé en 1993 ?

```
Dépendances : SUJ(se suicider, REPONSE)
              DATE(se suicider, 1993)
              attribut(REPONSE, ministre)
              attribut(ministre, premier)
```

Si on retrouve dans un texte la phrase *Pierre Bérégovoy s'est suicidé en 1993*, la variable REPONSE s'unifie avec *Pierre Bérégovoy* qui devient une réponse candidate : les deux premières dépendances de la question sont ainsi vérifiées dans cette phrase. Il manque les deux suivantes, concernant le type précis (Pierre Bérégovoy était-il premier ministre ?).

La validation est opérée en deux étapes. Tout d'abord, le système vérifie que la réponse candidate est bien un *ministre*, en recherchant une relation nommée 'attribut' (attribut(Bérégovoy, ministre)). Si cela est confirmé, le type étendu est également vérifié et les deux relations attribut(Bérégovoy, ministre) et attribut(français, ministre) sont attendues dans la même phrase.

### 4.4. Justification et classement des réponses candidates

#### 4.4.1. Classement des réponses

À chaque réponse est associé un quadruplet :

- la présence de tous les noms propres de la question dans la phrase (0 ou 1),

Véronique Moriceau, Xavier Tannier, Brigitte Grau.

- la réponse a le bon type d'entité nommée (0 ou 1),
- la réponse a le bon type précis (0, 1 pour le type simple, 2 pour le type étendu),
- le nombre de fois où la réponse a été extraite.

Les critères suivants sont utilisés pour classer les réponses (du plus important au moins important) :

- 1) Si le passage justificatif ne contient pas tous les noms propres de la question, alors la réponse est disqualifiée ;
- 2) Une réponse possédant le bon type d'entité nommée est préférée (quels que soit les scores ci-dessous) ;
- 3) Une réponse ayant validé le type précis est préférée également (le type étendu étant idéal) ;
- 4) Enfin, si tous les critères ci-dessus sont équivalents, une réponse trouvée plusieurs fois additionne les poids obtenus par chaque occurrence. La redondance est une information importante qui permet de masquer bon nombre d'erreurs.

#### 4.4.2. *Passage justificatif*

Le passage justificatif est produit de la façon suivante :

- Pour chaque réponse, la phrase ayant le meilleur score est sélectionnée.
- Les phrases précédentes sont également incluses dans le passage (dans le but de collecter le contexte et d'éventuels antécédents anaphoriques) dans une limite de 256 caractères (limite classique des campagnes d'évaluation).

## 5. Evaluation et résultats

Nous avons évalué notre système sur les données de test des campagnes d'évaluation CLEF 2005 et 2006 [VAL 05, MAG 06]. La collection de documents est composée d'environ 177000 articles de journaux en français (Le Monde et ATS 1994-1995 (environ 2 Go)). Ces documents sont censés être syntaxiquement corrects. Les questions de CLEF 2005 sont factuelles ou de type définition tandis que les questions de CLEF 2006 sont factuelles, de type définition ou liste.

Lors de ces campagnes, les systèmes sont autorisés à proposer 3 réponses pour chaque question, ces réponses devant être classées par ordre de confiance.

Plusieurs points doivent être évalués :

- les performances du système sur différents types de questions,
- l'apport de la stratégie de décomposition des questions dans le but de valider une réponse grâce à plusieurs passages.

Valider les réponses à des questions.

CLEF 2005		
Type de question	FIDJI	QRISTAL
Factuelle	<b>53%</b>	59%
Définition	<b>78%</b>	86%
<b>TOTAL</b>	<b>59.5%</b>	64%
CLEF 2006		
Type de question	FIDJI	QRISTAL
Factuelle	<b>46%</b>	64%
Définition	<b>56.5%</b>	83%
Liste	<b>50%</b>	50%
<b>TOTAL</b>	<b>48.5%</b>	68%

**Tableau 1.** Résultats de FIDJI sur les données de CLEF 2005 et 2006.

Position de la réponse correcte	Rang 1	Rang 1 à 3
<b>CLEF 2005</b>	59.5%	68.5%
<b>CLEF 2006</b>	48.5%	57%

**Tableau 2.** Position des réponses correctes.

### 5.1. Evaluation de FIDJI

Le tableau 1 présente les résultats obtenus par FIDJI sur les données de CLEF 2005 et 2006 (nombre de réponses correctes proposées en première position). Nous comparons ces résultats avec ceux de QRISTAL [LAU 05, LAU 06], le meilleur système de questions-réponses pour le français lors de ces campagnes.

Le tableau 2 précise le rang des réponses correctes proposées par FIDJI.

FIDJI obtient de moins bons résultats sur les données de CLEF 2006. Ceci est dû au fait que de nouveaux types de questions ont été introduits lors de CLEF 2006 : des questions de type définition et liste. En effet, pour les questions de type définition, de nouvelles catégories (par exemple, des définitions d'objets : *Qu'est-ce qu'un t-shirt ?*) ont été ajoutées afin de réduire le nombre de questions auxquelles il est facile de trouver une réponse (par exemple, les acronymes (*Qu'est-ce que RKA ?*) et les descriptions de personnes (*Qui est Bill Clinton ?*) sont souvent des appositions aux noms propres) [MAG 06]. Une autre difficulté se situe au niveau de l'identification des questions de type liste : des questions telles que *Citez le nom de tous les aéroports de Londres* sont facilement identifiées par les systèmes comme étant des questions de type liste alors que des questions telles que *Qui a découvert la comète Shoemaker-Levy ?* sont plus difficiles à analyser.

Véronique Moriceau, Xavier Tannier, Brigitte Grau.

Les résultats de FIDJI sont inférieurs à ceux de QRISTAL qui utilise aussi une approche syntaxique mais qui bénéficie de l'utilisation de nombreuses ressources telles que des dictionnaires et des ontologies. Cependant, notre système se place "virtuellement" à la deuxième place de ces campagnes d'évaluation puisque les systèmes qui ont atteint la deuxième place ont obtenu un score de 35% de réponses correctes pour CLEF2005 et de 46% pour CLEF2006.

## 5.2. Evaluation de la stratégie de décomposition des questions

Dans l'ensemble de questions de CLEF 2005, il y a 51 questions (25%) qui peuvent être décomposées en sous-questions en appliquant notre stratégie (à noter que seulement 17% des questions sont décomposables pour [HAR 08], qui utilise aussi une stratégie de décomposition similaire). FIDJI trouve une réponse correcte (sans tenir compte du rang) pour 32 questions. Chaque fois que la décomposition des questions est appliquée, le système peut rechercher des justifications à la réponse dans des documents différents.

Parmi les 32 réponses correctes pour les questions qui ont été décomposées en sous-questions, 22% d'entre elles ont une justification dans plusieurs documents. En revanche, si FIDJI n'utilise pas la stratégie de décomposition, seulement 64% des réponses sont correctes (au lieu de 68.5%).

Dans l'ensemble de questions de CLEF 2006, 56 questions peuvent être décomposées. FIDJI trouve une réponse correcte pour 29 d'entre elles : 14% de ces réponses correctes ont une justification dans plusieurs documents. Si FIDJI n'utilise pas la stratégie de décomposition, seulement 55% des réponses sont correctes (au lieu de 57%).

Le tableau 3 montre le nombre de réponses correctes dont la justification a été trouvée dans un ou plusieurs documents.

Parmi 32 réponses correctes	<b>CLEF 2005</b>
Justification dans 1 document	78%
Justification dans plusieurs documents	22%
Parmi 29 réponses correctes	<b>CLEF 2006</b>
Justification dans 1 document	86%
Justification dans plusieurs documents	14%

**Tableau 3.** Evaluation de la stratégie de décomposition des questions.

Ces résultats montrent une amélioration des performances du système lorsque l'on utilise une stratégie de décomposition des questions afin de valider les réponses par l'intermédiaire de plusieurs documents.

## 6. Conclusion

Dans cet article, nous avons présenté comment l'utilisation de la syntaxe peut aider un système de questions-réponses à produire de bons résultats. Les performances de notre système, évalué sur les collections CLEF, sont proches des meilleurs systèmes, et ceci sans avoir recours à de nombreuses ressources externes exploitées par ces systèmes. Les résultats montrent aussi que la validation de réponses par plusieurs documents améliore le traitement de certains types de questions. D'autres expérimentations sont en cours pour mesurer l'apport exact de l'analyse syntaxique dans chaque module du système.

Nous allons maintenant nous attacher à tester notre système sur une collection de documents provenant du Web dans le contexte du projet Quaero<sup>3</sup>. Un work package de Quaero est en effet consacré aux questions-réponses sur le Web. Un corpus de 2 millions de pages a été créé, et les questions sont définies et évaluées par un partenaire indépendant. Des questions de type liste, booléennes et complexes ('pourquoi', 'comment') seront proposées. Cette étude aura pour but d'étudier si des techniques de TAL peuvent être appliquées avec succès à de très grandes collections composées de documents de styles très différents. La taille de la collection de documents étant trop importante, nous modifions actuellement l'architecture afin d'éviter d'analyser entièrement le corpus.

Par ailleurs nous nous proposons aussi d'étudier la vérification de relations manquantes autres que le type de la réponse. Cela nécessitera sans doute de contextualiser la recherche afin de contrôler les inférences recherchées. Nous allons pour cela utiliser un nouvel ensemble de questions plus approprié à la recherche multi-documents.

## 7. Bibliographie

- [BOU 00] BOURIGAULT D., FABRE C., « Approche linguistique pour l'analyse syntaxique de corpus », *Cahiers de Grammaire*, vol. 25, 2000, Université Toulouse Le Mirail.
- [BOU 07] BOUMA G., FAHMI I., MUR J., VAN NOORD G., VAN DER PLAS L., TIEDEMANN J., « Linguistic knowledge and question answering », *Traitement automatique des langues*, vol. 46, 2007, Hermes-Lavoisier.
- [FOR 08] FORNER P., PEÑAS A., ALEGRIA I., FORASCU C., MOREAU N., OSENOVA P., PROKOPIDIS P., ROCHA P., SACALEANU B., SUTCLIFFE R., SANG E. T. K., Eds., *Working Notes for the CLEF 2008 Workshop*, Aarhus, Denmark, septembre 2008.
- [HAR 08] HARTRUMPF S., GLÖCKNER I., LEVELING J., « University of Hagen at QA@CLEF 2008 : Efficient Question Answering with Question Decomposition and Multiple Answer Streams », Forner et al. [FOR 08].
- [HAT 04] HATCHER E., GOSPODNETIĆ O., *Lucene in Action*, Manning, 2004.

---

3. <http://www.quaero.org>

Véronique Moriceau, Xavier Tannier, Brigitte Grau.

- [JAC 96] JACQUEMIN C., « A symbolic and surgical acquisition of terms through variation », *Connectionist, Statistical and Symbolic Approaches to Learning for Natural Language Processing*, Heidelberg, 1996.
- [KAT 03] KATZ B., LIN J., « Selectively using relations to improve precision in Question Answering », *Proceedings of workshop on Natural Language Processing for Question Answering, EACL*, Budapest, 2003.
- [KAT 05] KATZ B., BORCHARDT G., FELSHIN S., « Syntactic and Semantic decomposition Strategies for Question Answering from Multiple Resources », *Proceedings of the AAAI 2005 Workshop on Inference for Textual Question Answering*, Pittsburgh, 2005.
- [LAU 05] LAURENT D., SÉGUÉLA P., NÈGRE S., « Cross Lingual Question Answering using QRISTAL for CLEF 2005 », *Working Notes QA@CLEF*, Vienna, 2005.
- [LAU 06] LAURENT D., SÉGUÉLA P., NÈGRE S., « Cross Lingual Question Answering using QRISTAL for CLEF 2006 », *Working Notes QA@CLEF*, Alicante, 2006.
- [LIG 07] LIGOZAT A., « Apport de l'analyse syntaxique des phrases dans un système de questions-réponses », *Traitement automatique des langues*, vol. 46, 2007, Hermes-Lavoisier.
- [MAG 06] MAGNINI B., GIAMPICCOLO D., FORNER P., AYACHE C., OSENOVA P., PEÑAS A., JIKOUN V., SACALEANU B., ROCHA P., SUTCLIFFE R., « Overview of the CLEF 2006 Multilingual Question Answering Track », *Working Notes QA@CLEF*, Alicante, 2006.
- [MOR 08] MORICEAU V., TANNIER X., GRAPPY A., GRAU B., « Justification of Answers by Verification of Dependency Relations - The French AVE Task », Forner et al. [FOR 08].
- [ROD 08] RODRIGO A., PEÑAS A., VERDEJO F., « Overview of the Answer Validation Exercise 2008 », Forner et al. [FOR 08].
- [ROS 07] ROSSET S., GALIBERT O., ADDA G., BILINSKI E., « The LIMSI Qast systems : comparison between human and automatic rules generation for question-answering on speech transcriptions », *ASRU*, Kyoto, 2007.
- [VAL 05] VALLIN A., GIAMPICCOLO D., AUNIMO L., AYACHE C., OSENOVA P., PEÑAS A., DE RIJKE M., SACALEANU B., SANTOS D., SUTCLIFFE R., « Overview of the CLEF 2005 Multilingual Question Answering Track », *Working Notes QA@CLEF*, Vienna, 2005.