

Morphosyntaxe de l'interrogation pour le système de question-réponse RITEL

Anne Garcia-Fernandez (1), Carole Lailler (2)

(1) LIMSI-CNRS – Université Paris Sud,
BP 133 - 91403 Orsay Cedex
annegf@limsi.fr

(2) LIUM – Université du Maine,
72085 Le Mans Cedex 9
carole.lailler@lium.univ-lemans.fr

Résumé Nous proposons d'étudier le cas de l'interrogation en Dialogue Homme-Machine au sein d'un système de Question-Réponse à travers le prisme de la Grammaire Interactive. Celle-ci établit un rapport direct entre question et réponse et présuppose que la morphosyntaxe d'une interrogation dépend d'une « réponse escomptée »; l'interlocuteur humain ou machine ayant la possibilité de produire une réponse effective divergente. Nous proposons d'observer la présence des différentes formes de questions dans un corpus issu de l'utilisation du système RITEL. Et nous présentons une expérience menée sur des locuteurs natifs qui nous a permis de mettre en valeur la différence entre réponses effectives produites par nos sujets et réponses présupposées par le contenu intentionnel des questions. Les formalismes ainsi dégagés ont pour but de donner aux systèmes de DHM des fonctionnalités nouvelles comme la capacité à interpréter et à générer de la variabilité dans les énoncés produits.

Abstract In this paper we study the interrogation in Human-Computer Dialogue in order to integrate interpretative and dynamic functionalities in a Question-Answering system using Interactive Grammar. It constitutes a direct link between question and answer and assumes that the morphosyntax of interrogation depends on an anticipated answer. We propose an observation of question-answering corpora coming from the RITEL system showing the distribution of different morphosyntactic kinds of questions. Then we present an experimentation based on this observation in which French native speakers answer to those different kinds of questions. Our objective is to observe the difference between the anticipated answer and the effective one.

Mots-clés : Grammaire Interactive ; modèles dynamiques ; dialogue homme-machine ; système de question-réponse ; morphosyntaxe ; expérimentation utilisateur ; formulation de réponse ;

Keywords: Interactive Grammar ; dynamic models ; human-computer dialogue ; question-answering system ; morphosyntax ; user experimentation ; answer formulation ;

1 Introduction

Le point de départ de notre travail est d'élaborer un paradigme de l'interrogation en français moderne décliné en différentes versions selon le type de questions (quantification, localisation, temporalité, etc.). Nous avons travaillé à partir d'un corpus de Dialogue Homme-Machine (DHM) à travers le prisme de la Grammaire Interactive (GI) (Luzzati, 2006). En effet, d'une part, les interactions langagières dans le cadre de l'oral spontané ne se limitent pas au respect ou au non-respect d'une norme grammaticale mais sont le résultat d'énoncés perçus et conçus au fil de leur énonciation. D'autre part, bien que la théorie des actes de langage (Searle, 1969) autorise la planification des interactions, elle ne permet pas de lier contenu intentionnel des énoncés et phénomènes morphosyntaxiques. Le fait même d'établir un lien de dépendance entre réponse et question accrédite la validation d'un modèle grammatical par les productions elles-mêmes.

La formulation de la question étant conditionnée dans l'esprit du locuteur par une « réponse escomptée » et par l'orientation qu'il veut donner à son intervention, il s'agit moins d'appréhender un « objet question » capturé à l'instant t que d'explicitier, et à terme de faire fonctionner en DHM, un processus d'interaction langagière dynamique dans une utilisation réelle de la langue. La question est de savoir comment les variations morphosyntaxiques d'une question peuvent prédéterminer la forme de sa réponse. La morphosyntaxe fournit des éléments tangibles de réflexion. On peut donc légitimement s'y appuyer pour construire des modèles grammaticaux dynamiques implémentables qui permettent une réelle analyse automatique liant question et réponse. Les corpora d'énoncés de DHM téléphoniques issus de RITEL (Rosset, 2006) ont la particularité de présenter de nombreuses reformulations d'une même question et donc de comporter des échantillons de variation très large allant de la question sous sa forme la plus polie *Je voudrais savoir où se trouve Paris* à la question la plus agressive *Paris, c'est où?*

Les systèmes de question-réponse (SQR) ont pour objectif de permettre l'accès à une information en posant une question en langue naturelle à laquelle le système propose une réponse supposée elle aussi en langue naturelle. De nombreuses étapes relevant de la recherche d'informations sont nécessaires pour trouver l'information à retourner. Au-delà de fournir la bonne information, nous nous attachons ici au fait que cette réponse peut prendre plusieurs formes : elle peut être un mot, une phrase, un extrait de document, accompagnée ou non d'une justification,... Ici, nous nous intéressons au niveau phrastique, c'est-à-dire à la formulation de la réponse. Notre objectif est d'observer les réponses de locuteurs et de montrer le lien entre la forme de la réponse et la forme morphosyntaxique de la question afin d'utiliser cette dernière pour choisir la formulation de réponse la plus adaptée linguistiquement à la question.

Des travaux ont déjà montré l'importance de la forme syntaxique de la question dans la recherche de la réponse mais les applications sont majoritairement l'expansion de la requête et la recherche d'une réponse exacte. (Ligozat, 2005) montre que la forme syntaxique de la question permet d'élaborer de meilleurs patrons de recherche dans la phase de fouille de documents mais ne traite pas de la réalisation linguistique de la réponse. (Mendes, 2004) montre l'importance de prendre en compte la structure syntaxique de la question dans la réponse, mais s'appuie sur un domaine fermé. De plus, le choix de la formulation se fait *a priori*. Nous cherchons à nous distinguer de telles approches d'une part car RITEL se veut de domaine ouvert, d'autre part car la GI cherche, par essence, à prendre en compte la dynamique de l'échange dialogique. (Anaya, 2003) obtient des réponses de locuteurs natifs sous forme de longues phrases qui sont analysées syntaxiquement et classées. Une analyse similaire est appliquée aux questions. Après une analyse statistique, on détermine pour chaque forme syntaxique de question, la ou les formes syntaxiques de réponses possibles. Cependant, cette étude ne propose pas d'évaluation ou d'étude du lien entre question et réponse. Ici, nous avons

tenté, par un retour à la morphosyntaxe, d'expliciter ce lien. De récents travaux se sont centrés sur la formulation de la réponse, notamment dans des cas de réponses multiples ou peu sûres. Ce sont principalement les travaux présentés dans la thèse (Moriceau, 2007).

Dans un premier temps, nous nous efforcerons de présenter les modèles issus de la GI, notamment ceux des questions quantificatrices (QQ) en insistant sur leur caractère dynamique. Puis la présentation d'une étude expérimentale observant et analysant les réponses produites par des locuteurs face à l'ensemble des types morphosyntaxiques de questions définis par la GI, nous permettra de vérifier les liens entre formulation de la question et réponse effective.

2 Grammaire Interactive et réponses escomptées

La GI des questions permet de mieux appréhender le spectre de variabilité autour d'une formulation interrogative prototypique et ce selon deux axes. Nous présentons le cas des questions quantificatrices (QQ) qui sont des questions dont la réponse « escomptée » est une quantification (numérale ou non) par exemple *Combien coûte une baguette ?* Ces questions traitées dans (Luzzati, 2006) et sont ici reprises car le travail évoqué en troisième partie de cet article est fondé sur elles. Des versions dérivées pour les questions locatives, temporelles et en Entités Nommées ont été développées (Lailier, à paraître).

	1a adverbiale	1b numérale	1c déterminative	1d nominale	1e confirmative	Objet visé
2a prototypique	<i>Cb coûte cette voiture</i>	<i>Cb d'€ coûte cette voiture</i>	<i>Quel prix coûte cette voiture ?</i>	<i>Que coûte cette voiture ?</i>	<i>Cette voiture coûte-t-elle X € ?</i>	repérage quantitatif
2b tonique	<i>Cette voiture coûte cb ?</i>	<i>Cette voiture coûte cb d'€ ?</i>	<i>Cette voiture coûte quel prix ?</i>	<i>Cette voiture coûte quoi ?</i>	<i>Cette voiture coûte (bien) X € ?</i>	réalité du prix - alternative possible
2c renforcée	<i>Cb est-ce que coûte cette voiture ?</i>	<i>Cb d'€ est-ce que coûte cette voiture ?</i>	<i>Quel prix est-ce que coûte cette voiture ?</i>	<i>Qu'est-ce que coûte cette voiture ?</i>	<i>Est-ce que cette voiture coûte (bien) X € ?</i>	quantification
2d périphrastique	<i>Je voudrais savoir cb coûte cette voiture ?</i>	<i>Je voudrais savoir cb d'€ coûte cette voiture ?</i>	<i>Je voudrais savoir quel prix coûte cette voiture ?</i>	<i>Je voudrais savoir ce que coûte cette voiture ?</i>	<i>Je voudrais savoir si cette voiture coûte (bien) X € ?</i>	connaissance de la réponse
2e assertive	<i>Cette voiture coûte X € (?)</i>	<i>Cette voiture coûte X en € (?)</i>	<i>Le prix de cette voiture est (bien) de X €. (?)</i>	<i>Cette voiture coûte bien quelque chose ?</i>	<i>C'est bien X € que coûte cette voiture ?</i>	confirmation ou infirmation
Réponse	son prix est de X €, pas à vendre neutralité	X €, pourquoi, c'est 0 € numérisation	pas cher/cher, rien estimation	environ tant d'€, pas grand chose estimation	oui, pas du tout confirmation / infirmation	

Tableau 1 : Le modèle des QQ : questions dont la réponse est une quantification adapté de (Luzzati, 2006)

Nous sommes donc en présence de deux axes d'analyse (tableau 1). Le premier correspond à un inventaire ontologique des questions. (Gruber, 1993) parle ainsi « d'engagement ontologique minimal » dans la mesure où l'ontologie a pour but de définir un vocabulaire capable de circonscrire, en le décrivant, un domaine. Le second permet, quant à lui, de mesurer le degré de pertinence de telle ou telle partie de la GI selon son contexte d'utilisation. Le premier axe (en ligne dans le tableau), propre à l'expression du type mis en cause (la quantification dans notre cas, mais aussi la localisation, la temporalité,...), concerne la sémantique des questions. Il s'agit de faire varier de 1^a à 1^e l'objet de la question et la réalité

de la mise en cause dont va découler l'interrogation. La question prototypique correspond à une interrogation neutre, qui n'induit rien de particulier et attend une réponse simple, se limitant à l'information requise.

Le second axe (colonne), véritable *syntaxe de l'interrogation*, concerne la formulation des questions et ce quel que soit le mot interrogatif. Indépendamment du type même de la question mis en cause, il s'agit de varier les possibilités morphosyntaxiques de l'interrogation. De 2^a à 2^e, les outils de l'interrogation ainsi que l'intonation changent; ce qui a pour but de mettre l'accent sur les multiples possibilités d'effets de sens de la question. Sont ainsi mis en cause voire en défaut le locuteur ou son allocutaire (question tonique), leur maîtrise de la réponse (question périphrastique), le prix (question renforcée, qui est, nous semble-t-il, la formulation la plus ordinaire), la réalité du prix demandé ou une alternative sous-jacente (question confirmative). Il est nécessaire de ne pas perdre de vue que le mot interrogatif relève d'une classe grammaticale ou d'une autre suivant ce sur quoi le locuteur fait porter l'interrogation. En appréhendant plus précisément et plus sûrement le spectre de la variabilité, l'expression de l'interrogation en interaction sera davantage circonscrite tant morphosyntaxiquement que pragmatiquement. Certes, on accède par la GI à une intentionnalité qui n'est parfois qu'inconsciente. Néanmoins, cette grammaire a le mérite de réellement tenir compte de la variabilité du langage et, avec la génération, de l'exploiter. L'intérêt n'est pas mince puisqu'on peut espérer à terme, aboutir à des systèmes avec lesquels on acceptera de dialoguer.

L'axe ontologique de la quantification

Le français autorise l'expression de l'interrogation par des moyens morphologiquement différenciables où le mot interrogatif pointe sur la nature ontologique de la réponse escomptée. On peut donc qualifier une question par *combien* d'adverbiale (1^a). Cette dernière est dite prototypique et, de ce fait, véhicule un maximum de neutralité. Elle appelle un actant du verbe de sens quantitatif ou non. Une question utilisant le syntagme prépositionnel *à quel prix* ou *quel prix* est déterminative (1^c). Le locuteur fait, dans sa réponse escomptée, une estimation quant à la quantification. L'effet produit est ainsi moins neutre que dans le cas prototypique. Il ne s'agit pas de quantifier précisément l'objet visé mais bel et bien de circonscrire un périmètre où ce dernier doit se trouver. La question 1^b est dite numérale car elle demande une formulation chiffrée. Seules ces trois types de questions sont réellement quantitatifs. La question nominale (1^d) est, quant à elle, liée à l'existence d'un verbe quantitatif permettant l'adjonction d'un *que*. De même, avec inversion du sujet et/ou versation sur une question de type déterminatif, l'on obtient une question propositionnelle (ou une interrogation totale) qui ne cherche que la confirmation ou l'infirmité. Celle-ci ne devient quantitative qu'en présence d'un verbe à valeur quantitative ou d'un possible contexte quantitatif accompagné, par exemple, d'un adverbe locatif évaluatif *cher/pas cher*.

Les différentes questions ont une portée spécifique, dans la mesure où elles concernent le prix en lui-même, une estimation de ce prix, sa réalité ou bien sa légitimité. Les réponses oscillent entre une réponse positive qui va affirmer l'existence tangible du prix et d'une réponse négative qui va présenter diverses façons de prendre parti entre la gratuité et le refus d'apprécier le prix (l'estimation, l'affirmation de l'erreur, l'infirmité). Le recours à de telles formulations (différentes bien qu'appartenant au même spectre) présuppose bel et bien que le locuteur fait varier le statut de sa question et qu'il escompte, sinon des réponses différentes, du moins des façons différentes de répondre à sa question.

L'axe de l'interrogation

Parallèlement à la question qu'il souhaite poser, tout locuteur est en mesure de produire une formulation de cette même question plus ou moins renforcée, et ce de différentes manières. Il peut choisir de postposer l'outil interrogatif ; ce dernier devenant tonique (2^b). Ainsi, la question *Cette baguette coûte combien ?* qui est une interrogation tonique « classique »

suppose un certain degré d'agressivité. Le locuteur est pour le moins agacé de l'absence de réponse ou de la volonté de son allocutaire de ne pas répondre. De plus, la question tonique *Elle coûte combien cette baguette ?* induit un niveau supplémentaire d'agressivité. Le locuteur, par ce procédé de clivage, peut, par exemple, sous-entendre que le prix de la baguette n'est pas légitime car trop élevé. L'allocutaire, lui aussi accusé, ressent pleinement le degré d'agressivité du *questionneur*. Les faces de chacun (voir par exemple (Goffman, 1974) et (Kerbrat-Orecchioni, 2005)) ne sont plus préservées et la conversation risque d'être interrompue ou parasitée.

D'autres formes de renforcement sont envisageables. La locution interrogative renforcée *est-ce que* (2^c), l'interrogation périphrastique ou l'interrogation indirecte du type *Je voudrais savoir (si)* (2^d) constituent autant de formes renforcées. Ce dernier énoncé introducteur efface toutefois toute trace morphosyntaxique de l'interrogation.

Les interrogations simples (2^a), quant à elles, engendrent – si le locuteur souhaite répondre pleinement aux critères de normativité– l'inversion du sujet *Combien coûte cette baguette ?* ou la version *Combien cette voiture coûte-t-elle ?* et n'acceptent que difficilement la postposition sauf cas de focalisation *C'est combien qu'elle coûte cette voiture ?*

Enfin, l'interrogation assertive (2^e), marquée à l'oral par l'intonation, ne permet l'insertion que dans le cadre du clivage *Elle coûte bien X euros, cette voiture ?* et un marqueur de tonicité vient très souvent renforcer l'assertion.

Il semble que la seule véritable interrogation soit la question renforcée car elle seule a pour unique objet visé le prix et non son estimation ou sa légitimité. L'interrogation périphrastique est celle, qui de toutes, est la plus polie et celle qui ménage le plus l'allocutaire. En ne présupposant pas qu'il doit connaître la réponse, on ménage sa face négative, y compris s'il ne la connaît effectivement pas. C'est d'ailleurs la question la plus fréquente au sein de RITEL car le locuteur, soucieux de se faire comprendre de la machine, s'efforce de poser une question sous sa forme la plus claire et la plus polie.

En confrontant le modèle de la GI des QQ à un corpus d'énoncés de DHM et en analysant les modalités de catégorisation des questions, nous pourrions également mesurer le degré de pertinence du modèle et, le cas échéant, le faire évoluer en prenant notamment en compte les réponses humaines apportées.

	1a	1b	1c	1d	1e	Total
2a	4	32	14	0	0	50 (15,9%)
2b	0	2	0	0	0	2 (0,6%)
2c	0	5	0	0	1	9 (2,9%)
2d	5	146	97	0	2	250 (79,6%)
2e	0	3	0	0		3 (1,0%)
Total	9 (2,9%)	185 (58,9%)	117 (37,2%)	0	3 (1,0%)	314

Tableau 2 : Les résultats – présence de QQ dans RITEL tiré de (Luzzati, 2006)

Le corpus RITEL est un corpus de requêtes utilisateurs du LIMSI (Galibert, 2005) (Rosset, 2006) en situation de Dialogue Oral Homme Machine (DOHM) collecté de septembre 2004 à janvier 2005 dans lequel 13 personnes qui ont reçu, à titre d'exemple, une liste différente d'environ 300 questions ont appelé un serveur, ce qui a donné lieu à 582 dialogues qui comportent 5362 énoncés. Les locuteurs obéissaient à une consigne simple : « le système a beaucoup de mal à reconnaître et à comprendre ce que vous lui dites, mais il a besoin de

données pour apprendre : essayez d'obtenir des réponses et de lui expliquer ce que vous cherchez ». Il en résulte un large corpus d'énoncés très divers que balaye de temps à autre le mot interrogatif *combien* (cf. tableau 2). Les QQ se répartissent principalement en 314 interrogations qui constituent soit la réitération de plus en plus insistante d'une même question, soit sa modulation, soit enfin l'intégration d'une QQ au sein d'une thématique.

Les 314 QQ du corpus RITEL sont toutes catégorisables à l'aide des critères de la GI. Pour quelques occurrences, des hésitations sont possibles quant à la catégorisation à mettre en œuvre (Luzzati, 2006) mais il est notable de constater que cette grammaire fonctionne comme un outil d'analyse de *corpora* qui n'ont pas été construits pour elle.

Le concept de *réponse escomptée* permet donc de postuler qu'il existe une pré-réponse ou du moins une image mentale de cette réponse en fonction de laquelle se constitue la question posée. Chaque locuteur aurait ainsi en sa possession un spectre de variabilité qu'il pourrait faire jouer en fonction de l'allocutaire, de ce qu'il sous-entend (consciemment ou non) et/ou de ce qu'il cherche à obtenir comme information.

3 Observation de réponses effectives par des locuteurs natifs

La GI comme nous l'avons vu permet d'appréhender, selon la morphosyntaxe de la question, l'intentionnalité du locuteur et ce qu'il met en cause. Cependant, lorsqu'on pose une question tonique numérale par exemple, il se peut que la réponse soit toute autre que la réponse escomptée. Par exemple, à la question *Combien est-ce qu'une baguette de pain coûte ?*, une réponse escomptée est : *Une baguette coûte (environ) 80 centimes d'euros* mais des réponses possibles sont : *Une baguette coûte environ 80 centimes; Une baguette peut coûter de 75 centimes à 1 euro; Je ne sais pas, moins d'un euro; Ça dépend, tu vas à quelle boulangerie; Tu devrais le savoir;...* C'est le contexte et la volonté du locuteur répondant qui jouent. Connaît-il ou non la réponse ? A-t-il envie de répondre à la question ? A-t-il besoin de plus de précision ? L'expérimentation proposée ici permet l'observation des réponses effectives face aux différentes formes morphosyntaxiques des questions décrites par la GI.

Nous proposons de comparer notre système de question-réponse à un utilisateur qui *veut* répondre et va vraiment aller à la recherche de la réponse. En raison des problèmes - notamment de reconnaissance vocale - qui peuvent se poser, l'utilisateur cherchera à confirmer les éléments identifiés dans la question au cours de sa réponse. L'expérimentation cherche donc à pousser nos sujets à répondre sous forme de phrases plutôt que par l'*information réponse* brute. Nous avons ainsi fait appel à des locuteurs natifs du français en leur demandant de répondre à des questions supposées posées par des enfants de CM2 préparant des exposés sur différents thèmes. Selon cette même idée, les réponses aux questions sont supposées faciles à trouver. Par ailleurs, cette étude concerne spécifiquement les QQ et se déroule à l'écrit (au travers d'une interface web). Des études suivront permettant notamment d'observer les réponses à l'oral mais aussi de comparer les réponses à une même question par un même sujet, à l'oral comme à l'écrit.

L'objet visé (cf. tableau 1) joue un rôle sur la formulation de la réponse. Deux facteurs ont été pris en compte. D'une part, la quantification de l'objet de la question est-elle fixe ou bien varie-t-elle (en fonction d'une variabilité interne ou bien d'un paramètre externe) ? En fonction de cet élément, on recherche un *type d'information* différent : *fixe*, *variable* ou dépendant d'un *paramètre* externe. D'autre part, nous supposons que le *type de mesure* concerné joue un rôle, en d'autres termes, nous supposons qu'on n'appréhende pas de la même façon des poids, des distances, des températures, etc. Le but de l'étude n'est pas d'évaluer l'impact de ces deux facteurs. Nous avons cependant constitué le corpus de questions tests en balayant différentes

modalités de ceux-ci afin d'éviter un biais face au choix d'un unique *type de mesure* ou encore d'un unique *type d'information*. Le tableau 3 présente les 9 formes canoniques que nous avons fait varier selon les axes de la quantification et de l'interrogation proposée par la GI. Nous avons donc un total de 207¹ *questions test*.

type de mesure type d'info.	Poids	Espace/taille	Temps/durée
Fixe	<i>Combien pèse une brique de lait ?</i>	<i>Combien mesure un double décimètre ?</i>	<i>Combien dure un mandat présidentiel en France ?</i>
Variable (interne)	<i>Combien pèse un bébé à la naissance ?</i>	<i>Combien mesure un bébé à la naissance ?</i>	<i>Combien dure une grossesse ?</i>
Paramètre (externe)	<i>Combien pèse une bouteille d'eau ?</i>	<i>Combien mesure une feuille A4 ?</i>	<i>Combien dure un mois de février ?</i>

Tableau 3 : Questions canoniques utilisées pour l'expérimentation

La passation des tests a été prévue pour permettre une comparaison d'au minimum 3 réponses par *question test* c'est-à-dire pour une même forme morphosyntaxique, un même type de mesure et un même type d'information. On a donc un minimum de 27 réponses comparables pour un même type morphosyntaxique ; chaque sujet étant face à un total de 18 questions.

Observation des résultats

L'analyse des réponses s'est faite selon une série de variables d'observation (explicitées au fur et à mesure par la suite et listées dans le tableau 4) qui en permettent une vue synthétique. Ces observations permettent de décrire l'ensemble des réponses formulées par les utilisateurs. Aucune annotation n'exclut les autres : elles sont complémentaires. Ce corpus de réponses (le *corpusR*) est constitué de 594 réponses. Nous tenons à souligner que les variables d'observation ont été élaborées uniquement au regard du *corpusR*, sans accès ni aux questions, ni aux réponses escomptées décrites par la GI et ce afin d'éviter tout biais d'annotation.

Observations :	fermée	complétive	doute	estimation	suggestion	reprise	expression facteur	hors sujet	valence positive ²
% du corpus annoté positivement :	23%	14,5%	5%	16%	3,8%	19%	33,9%	2%	95%

Tableau 4 : Pourcentage d'annotations positives par variable d'observation

Les réponses fermées

La variable *réponses fermées* concerne les réponses par *oui*, *non*, *effectivement*, *pas du tout*, etc. Ce ne sont presque exclusivement que les questions 1^e (questions propositionnelles - 80% d'entre elles) et 2^e (assertives - 78% d'entre elles) qui ont obtenu des réponses fermées (on a une occurrence sur 138 pour une question 1^{d2c}). 68% des réponses fermées sont constituées de cette réponse fermée seule. La liste des vocables utilisés est : *oui*, *non*, *cela dépend*, *c'est plausible*, *possible*, *exactement*, *en effet*. On a une seule occurrence de réponse confirmatrice

¹ 9 x les 23 formes de questions

² Etant donné la nature des questions, nous ne nous attendions pas à un grand nombre de réponses à valence négative ce qui est le cas (5%).

(en effet pour la question 1^{d2c}) type de réponse escomptée pour les 2^e. Les réponses escomptées pour les 1^e sont bel et bien celles choisies par nos sujets. Ce n'est pas le cas des questions 2^e.

Les réponses complétives

Une complétion est définie comme une réponse anticipant les besoins de l'utilisateur. Par exemple : *Oui. Avant ça durait 7 ans mais ça a changé.* C'est notamment le cas, des réponses explicatives et celles proposant des informations de calcul. On observe un taux de réponses complétives de 14,5%. N'ayant pas de corpus comparable, il est difficile d'évaluer l'importance de ce pourcentage. Cependant, ces 14,5% représentent près de 90 occurrences de réponses complétives qui pourront être étudiées et utilisées au sein de RITEL. Aucune réponse escomptée n'est de nature complétive. En fait, on peut supposer que la volonté de compléter la réponse est propre au *répondant* et n'est pas une attente particulière exprimée par le *questionnant*.

Mises en cause : expression du doute et estimations

On trouve dans le *corpusR* 28 réponses contenant une expression du doute. Il s'agit d'expressions telles que *il me semble que, je dirais, je sais pas trop*, etc. Bien que les questions 2^d (périphrastiques) présupposent un doute sur la maîtrise de l'information et devraient donc laisser plus de place à l'interlocuteur quant au fait de ne pas savoir ou d'avoir un doute sur sa propre connaissance, on a seulement 5% des 2^d qui expriment un doute. La répartition (tableau 5) montre cependant que les expressions de doute ne correspondent pas particulièrement à des questions 2^d.

1 ^a	1 ^b	1 ^c	1 ^d	1 ^e	2 ^a	2 ^b	2 ^c	2 ^d	2 ^e
25%	32%	21%	18%	4%	18%	28%	11%	25%	18%

Tableau 5 : Répartition des expressions de doute selon les axes horizontaux et verticaux

Les estimations sont une autre forme de doute qui concerne l'exactitude de la réponse et non pas la maîtrise de l'information. Elles couvrent 16% du corpus total et s'illustrent par des expressions telles que *environ, en moyenne, généralement*, etc. Tout comme l'expression du doute, on n'observe pas de différence selon les deux axes de la GI. En revanche, les répartitions diffèrent en fonction du type d'information et de mesure.

Il semble donc, concernant la mise en cause, qu'il n'y a pas d'influence de la forme morphosyntaxique de la question sur la réponse produite par nos sujets.

Suggestions

Les réponses suggestives (22 occurrences) sont définies comme contenant une aide à l'utilisateur dans l'expression de son besoin d'information. Dans le *corpusR*, ces cas correspondent à 91% à des *corrections* (*Que mesure un bébé à la naissance ? Tu veux dire "Combien mesure un bébé à la naissance?", non?*) c'est-à-dire une réponse par laquelle le sujet va exprimer que la question a été mal formulée ou bien est ambiguë. Il est à noter que nous n'envisageons pas de proposer de telles corrections à nos utilisateurs. D'une part, le but d'un système de question-réponse est différent, d'autre part les connaissances linguistiques du système RITEL sont insuffisantes.

Dans l'ensemble, très peu de réponses sont suggestives. 60% d'entre elles sont des 1^d et 25% sont des 1^{d2a}. On peut donc penser que de telles formulations sont considérées comme *maladroites* ou *imprécises* par les utilisateurs. Au cours de l'annotation du corpus RITEL selon la GI des QQ, aucune occurrence de ce type de question n'avait d'ailleurs été relevée.

Reprises

Les cas de reprise sont plus intéressants : parce que le système commet des erreurs (notamment de reconnaissance vocale), il est primordial, dans la réponse, de reformuler les éléments compris de la question. Donnant ainsi à voir à l'utilisateur les erreurs possibles, on confirme ces éléments et on offre l'opportunité à l'utilisateur de les rectifier. Au sein de notre corpus, 1/5 des réponses reprennent la question. Ce taux semble assez faible compte tenu du fait que, pour RITEL, il faudrait une reformulation systématique. Quoi qu'il en soit, nous disposons de 114 occurrences sur lesquelles nous pourrions prendre exemple pour formuler nos reprises.

Expression d'un critère et nombre de valeur numérique

L'expression d'un critère (qu'il soit restrictif ou sélectif) se réalise de différentes manières dans le *corpusR* : par une condition (*s'il arrive à terme*), par une énumération (*5 ans depuis 2002, avant il durait 7 ans*), simplement en précisant le facteur (*28 ou 29 jours, suivant les années*), en indiquant la valeur choisie pour ce facteur (*maintenant, oui*), etc. 1 réponse sur 3 exprime un critère. Ce résultat est très intéressant car il montre la volonté de la part des sujets (peut-être due au contexte : il faut répondre à des enfants) d'être précis dans leur réponse. On n'observe aucune prédominance de l'expression d'un facteur en fonction d'un type de question donné.

En observant le nombre de *valeur numérique réponse*³, on observe que dans 80% des cas d'expression d'un critère, il y a plus d'une valeur numérique répondue alors que c'est le cas pour seulement 16% des réponses sans expression d'un critère. Les sujets, plutôt que de faire un choix, préfèrent l'exhaustivité et donnent une valeur réponse pour chaque valeur du critère qu'ils identifient. Ce résultat amène à penser que donner une réponse unique n'est pas la meilleure solution et ce quelle que soit la forme de la question.

Autres

Afin de décrire totalement les réponses, nous avons utilisé la variable *hors sujet* pour les réponses contenant du hors sujet (par exemple : *Souvent trop longtemps. En ce moment, c'est 5 ans.*). Elles ne concernent que quelques cas isolés.

Applications en question-réponse

Le seul lien de corrélation effectif entre forme morphosyntaxique de la question et réponse produite se situe donc au niveau des réponses fermées. Nous ne pouvons donc pas utiliser directement les résultats de cette expérimentation au sein de notre système. Cependant, nous avons observé de nombreux faits qui permettront d'orienter la génération des réponses au sein d'un SQR tel que RITEL. D'une part les formulations de réponses obtenues nous serviront à établir des patrons morphosyntaxiques de réponses pouvant être directement utilisés par notre module de génération en langue. D'autre part nous conservons le but de formuler une réponse en fonction du type de la question. Nous envisageons ainsi d'extraire des paires *type de question - type de réponses* plus récurrentes que d'autres. Ces types de question, s'ils ne correspondent pas à une forme morphosyntaxique commune, serviront au moins à choisir une formulation de la réponse plus adaptée et nous permettront peut-être d'observer des éléments de la question qui jouent plus concrètement un rôle dans la formulation de la réponse par nos sujets. Enfin nous disposons grâce à cette expérimentation d'un corpus QR et d'un corpus de réponses contenant de nombreuses paraphrases. L'avantage étant que toutes les réponses sont avérées.

³ Ce chiffre ne prend pas en compte le « 2 » dans, par exemple « Oui, si la bouteille fait 2 litres » et a été pondéré afin de ne considérer les intervalles (donc deux valeurs numériques) comme une unique *valeur numérique réponse*.

4 Conclusions et perspectives

Nous avons pu montrer que la GI est efficace pour décrire le lien dynamique entre question et réponse, et ce en s'appuyant sur la morphosyntaxe des questions. L'intentionnalité qu'elle véhicule définit un ensemble de réponses escomptées. Au sein d'un SQR, cette approche a pour but d'éviter une interaction figée et rébarbative, véritable négation du dialogue.

Un apport majeur de l'expérimentation est l'obtention d'un corpus avéré de Q-R à l'écrit. Nous souhaitons à présent pouvoir comparer ce corpus à un corpus oral obtenu dans les mêmes conditions. Cette expérimentation a cependant pu montrer un écart certain entre réponse escomptée et réponse observée. Le modèle ne peut donc pas être utilisé de façon prédictive. Reste toutefois à savoir si les réponses escomptées qu'il propose ne produisent pas chez le questionnant davantage de satisfaction. Une nouvelle expérience de mesure de satisfaction face à différents types de réponses pour une même question dont une réponse escomptée est maintenant à envisager.

Références

- ANAYA G., KOSSEIM L. (2003). Generation of natural responses through syntactic patterns. Actes de *TALN Workshop Question-Réponse*, 2003.
- GALIBERT O., ILLOUZ G., ROSSET S., (2005). Ritel : an Open-Domain, Human-Computer Dialog System. Actes d'*Eurospeech'05*.
- GOFFMAN E. (1974). *Les rites d'interaction*, Paris : Edition de minuit.
- GRUBER TH. R. (1993). *Formal ontology in conceptual analysis and knowledge representation*, Kluwer academic Publishers.
- KERBAT-ORECCHIONI K. (2005). *Le discours en interaction*, Paris : Armand Colin.
- LIGOZAT A. (2005). Apport de l'analyse syntaxique des phrases dans un système de questions-réponses. *Traitement Automatique des Langues*, vol. 46, n°3, 103-125.
- LUZZATI D. (2006). Essai de description interactive : l'exemple des questions quantificatrices. Colloque *La quantification*, 15p. A paraître. Consultable sur le site de l'auteur : <http://www-ic2.univ-lemans.fr/luzzati>.
- MENDES S., MORICEAU V. (2004). L'analyse des questions: intérêts pour la génération des réponses. Actes de *TALN Workshop Question-Réponse*, 2004.
- MORICEAU V. (2007). *Intégration de données dans un système question-réponse sur le Web*. Thèse de doctorat en Informatique de l'Université Paul Sabatier - Toulouse III.
- ROSSET S., PETEL S. (2006). The Ritel Corpus - An annotated Human-Machine open-domain question answering spoken dialog corpus. Actes de *LREC'06*.
- ROSSET S., GALIBERT O., ILLOUZ G., MAX A. (2007) Interaction et recherche d'information : le projet RITEL, de la langue française. Actes de *TAL*, vol. 46 n°3, 156-179.
- SEARLE J. R. (1969). *Les Actes de Langage*, essai de philosophie du langage Paris : Hermann.