MACAQ : A Multi Annotated Corpus to study how we adapt Answers to various Questions

Anne Garcia-Fernandez, Sophie Rosset, Anne Vilnat

LIMSI - CNRS
F-91403 Orsay Cedex
{annegf, rosset, vilnat}@limsi.fr

Abstract

This paper presents a new corpus of human answers in natural language. The answers were collected in order to build a base of examples useful when generating natural language answers. We present the corpus and the approach we used for its acquisition. Answers correspond to questions with fixed linguistic form, focus, and topic. Answers to a given question exist for two modalities of interaction: oral and written. The whole corpus of answers was annotated both manually and automatically on different levels including for the most innovative: words from the questions being reused in the answer, the precise sentence part answering the question, which we define "answering-information", completions. A detailed description of each annotation is presented. Two examples of corpus analyses are described. The first analysis shows some differences between oral and written modality especially in terms of length of the answers. The second analysis concerns the reuse of the question focus in the answers.

1. Introduction

This paper presents a corpus of human answers in natural language collected in order to build a base of examples useful when generating natural language answers.

Question-answering (QA) is the task of automatically answering a question asked in natural language. From a question and a set of documents, question-answering systems extract and provide an answer. Most of these systems extract the information which answers the question from a single document and return it without including it in a sentence (Figure 1). Typically, QA systems return a minimal answer and a justification (extract of the document(s) from which the answer was extracted).

Question: *Where is the Mona Lisa?*
Answer: *Louvre*
Extract: *Mona Lisa (also known as La Gioconda) is a 16th century portrait painted in oil on a poplar panel by Leonardo da Vinci during the Italian Renaissance. The work is owned by the Government of France and is on the wall in the Louvre in Paris, France with the title Portrait of Lisa Gherardini, wife of Francesco del Giocondo.*

Figure 1: Example of question, answer and extract

Recently however, a number of systems have proposed to manage interactive QA (TREC ciQA task, 2007). Regarding interactions with a virtual agent or human-machine dialogues for instance, we assume that such an interaction requires answers in natural language rather than an extract of a document.

Since we work on the open-domain QA system RITEL(Toney et al., 2008), we cannot afford to build lists of patterns or canned texts (McDonald, 2003).

To generate answers in natural language, we choose to observe how human answers are formulated and, from those observations, create an answer generation model. Thus we collected a corpus of human answers in natural language. Our approach consists of two steps. We first manually generated a corpus of French questions with a fixed linguistic form (Garcia-Fernandez et al., 2009). Then we collected the corresponding answers from native French speakers. The collection was done in both speech and written modality and a transcription of spoken answers was carried out so the resulting corpus contains written, oral, and transcribed answers.

In order to compare answers (depending on the modality, on the question features, etc.) we needed a precise description for them. We proceeded to multi-level automatic annotations (part-of-speech tagging, syntactic analyses, etc.) and a manual annotation (on semantic and pragmatic levels). Other numerical features were computed, such as the length in words of the answer or the number of information-answers[1] in the answer.

We detail the corpus acquisition method and the answers corpus in sections 2 and 3. Section 4 presents a general description of the answers corpus and section 5 details different annotations of the corpus and how they could be used. Section 6 proposes two analyses as examples of how to exploit the corpus.

2. Corpus acquisition methodology

To observe human answers, we set up an experiment. Here the system does not answer questions asked by users (or given in a file as in evaluation campaigns). Instead people were asked to answer a set of questions proposed by the system. This protocol is unique. Although related work observes human answers, none of them allow an observation of several modalities (speech and written) for a common set of questions, keeping control on the syntactic and semantic

[1] In the extract of the example 1, "Louvre" is an information-answer.

structure of the question. Moreover our panel of subjects is larger than most others: 40 for (**?**), 152 in our case.

As we want to observe how the answer is formulated and presented, we proposed a context encouraging the subjects to compose complete sentences including the answer and not just words or short answers. We asked *easy* questions (about quantity, location or time and about general culture knowledge) hoping to minimize negative valence answers (such as "I don't know").

This context had to fit with the easiness of the questions, thus we asked native French speakers to answer questions supposedly asked by 10-years-old children preparing a poster at school. This context is particularly interesting because are naturally incitated to answer entire sentences.

Two platforms were used to collect data. For the written modality, a web site proposed a set of questions and corresponding text areas of few lines reserved for the user answers. For the speech modality, we used the existing RITEL platform (Toney et al., 2008): phone lines, speech detection system (detecting when the user starts and ends talking), speech synthesis (a unique vocal model for all tests). For both modalities, the same experimental context and number of questions were used.

The experiment consisted of two phases. The first one concerned a restricted set of questions (quantity questions) on both modalities. Each subject was asked 18 questions. After this first phase, we asked participants to give feed-back on the experiment. Thanks to this, we decided to increase the number of questions. Thus in the second phase, we extended the corpus to time and location questions and asked 24 questions to each subject. We contacted more than 1100 people,[2] among whom 203 accepted to participate (18.5% of contacted people). After rejecting all failure situations (the person accepted but was not a French native speaker, the person received all information but did not do the experiment, a problem occurred during the experiment,...) we had 152 participants (13 % of contacted people).

3. Corpus of questions

Questions are factoid and simple. They consist of quantity, time or location questions. Question topics are chosen to be easy to answer (French general knowledge). Moreover we took the nature of the answer into account: either there is one unique answer, or there are more than one possible answer. Most of the questions are composed by the minimal set: question markers, one principal verb and a focus defined as the nominal group representing the unit on which information is requested (Ferret et al., 2002). We added information to some of them to avoid ambiguity or to make the question more precise.

From a small set of basic questions (19), we generated 507 linguistic variations (examples will be given in following subsections 6.). It is a way to avoid having always the same structure of question and so have an experiment which is less boring for our participants. On the other hand, we wanted to have the possibility to compare answers with each other, depending on the linguistic form of the question.

(Luzzati, 2006) has recently proposed a model for question answering in interaction. It shows that the formulation of a question expresses the intention of the locutor and can thus be an indication of the linguistic form of the expected answer. We are not assuming that there is a unique correspondence between one question form proposed by the model and one answer form. But, we use this model for two reasons: (1) it proposes a set of morphosyntactic variations from a prototypical question and (2) it can be used as baseline establishing links between question and answer forms. Thus, for each semantic type, different syntactic forms are built.

For each question, we fixed the following features: *semantic type, semantic sub-type, syntactic form of the interrogative, syntactic form of the question,* and *lexical choices*. For each question, information on expected answers is also fixed: its *general type* and its *nature*. Following subsections detail these features.

3.1. Semantic type of the question

The corpus of questions is composed of time, location, and quantity questions. Table 1 shows an example for each type.

Semantic type	Example
Quantity	Combien pèse une bouteille d'eau ? *How heavy is a bottle of water?*
Location	Où est la Joconde ? *Where is the Mona Lisa?*
Time	Quand sont les Jeux Olympiques ? *When are the Olympic Games?*

Table 1: Question semantic type

3.2. Semantic sub-type of the question

For quantity questions, three semantic sub-types were tested. Table 2 gives examples.

Semantic subtype	Example
Weight	Combien pèse un bébé ? *How heavy is a baby?*
Duration	Combien dure une grossesse ? *How long is a pregnancy?*
Distance	Combien mesure un bébé ? *How tall is a baby?*

Table 2: Quantity question semantic subtype

3.3. Interrogative forms

Questions are built using different interrogatives. Table 3 shows examples for a location question about the Mona Lisa.

For quantity questions, two other interrogative forms are possible. Table 4 shows examples for a quantity question about the size of a baby.

[2] University students, friends, colleagues and people from the "faites-la-science" list (`http://faites-la-science.risc.cnrs.fr/`)

Syntactic form	Example	
Prototypical	Où est la Joconde ?	*Where is the Mona Lisa?*
Assertive	La Joconde est au Louvre ?	*Is the Mona Lisa in the Louvre Museum?*
Periphrastic	Je voudrais savoir où est la Joconde	*I would like to know where the Mona Lisa is?*
Reinforced	Où est-ce que se trouve la Joconde ?	*Where can it be found, the Mona Lisa?*
Tonic	La Joconde se trouve où ?	*The Mona Lisa is where?*

Table 5: Question syntactic forms

Interrogative form	Example
Adverbial	Où est la Joconde ?
	Where is the Mona Lisa?
Confirmative	La Joconde se trouve-t-elle au Louvre ?
	Is the Mona Lisa in the Louvre Museum?
Determinative	Dans quel musée se trouve la Joconde ?
	In which museum is the Mona Lisa?

Table 3: Interrogative forms

Interrogative from	Example
Nominal	Que mesure un bébé ?
	What does a baby measure?
Numeral	Combien de centimètres mesure un bébé ?
	How many centimeters does a baby measure?

Table 4: Quantity questions with specific interrogative forms

3.4. Question syntactic form

Different syntactic structures can be used in French to formulate the same question. Table 5 shows examples of different syntactic forms for a location question about the Mona Lisa.

3.5. Lexical choice: the verb

For time and location questions, the same variation of question appears twice: with a verb specific to the question semantic type (verb of location or time) or with a neutral verb (auxiliary). Table 6 shows a pair of question examples.

Verb type	Example
Auxiliary	Où est la Joconde ?
	Where is the Mona Lisa?
Location verb	Où se trouve la Joconde ?
	Where is the Mona Lisa located?

Table 6: Verb type

3.6. Expected answer type

A question expects a given type of answer. It can be a named entity (“location”, “time”, or “number” for time, location, and quantity questions) or a closed answer (as “yes”, “no”,...) in the case of closed questions. Table 7 shows expected answer types for questions about the Mona Lisa.

Answer type	Example
Yes-No answer	La Joconde se trouve au Louvre ?
	Is the Mona Lisa in the Louvre Museum?
NE country	Dans quel pays est la Joconde ?
	In which country is the Mona Lisa?
NE museum	Dans quel musée se trouve la Joconde ?
	In which museum is the Mona Lisa?
NE unknown	Où est la Joconde ?
	Where is the Mona Lisa?

Table 7: Answer type (with *NE for Named Entity*)

3.7. Answer nature

Depending on the object of the question, the answer could be fixed, or variable. Table 8 shows examples of questions for each answer nature. In the first example, the size of an A4 paper sheet is *fixed:* there is one unique answer. On the other hand, the duration of February depends on the year considered and so the answer is considered as *variable*.

Answer nature	Example
Fixed	Combien mesure une feuille A4 ?
	What size is an A4 paper sheet?
Variable	Combien dure février ?
	How long is February?

Table 8: Nature of the answer

4. General description of the answers corpus

Table 9 presents the characteristics of the entire final corpus given the modality axis.

	Written	Speech	Total
# answers	2,088	1,044	3,132
# different questions	507	493	507
# subjects	99	53	152
# subjects/question	4.12	2.12	6.17
# words	17,976	7,128	25,104
# different words	3,363	1,634	4,574
avg words/answer	8.39	5.99	7.61
avg duration (sec)/answer	33.0	4.2	21.0

Table 9: General characteristics of the corpus

The final corpus consists of 3,132 answers, among which 2,088 are written and 1,044 are spoken answers. In average

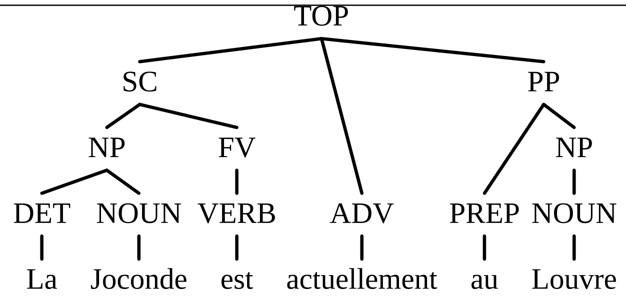
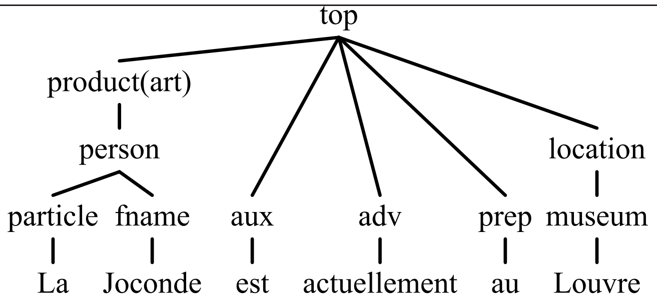
Version	Example					
Raw	La The	Joconde Mona Lisa	est is	actuellement currently	au in the	Louvre Louvre museum
Lemmatised	le	Joconde	être	actuellement	au	Louvre
POS	DET	NAM	VER	ADV	PRP:det	NAM
Syntactic Parsing	 TOP SC PP NP FV NP DET NOUN VERB ADV PREP NOUN La Joconde est actuellement au Louvre					
NCA	 top product(art) person location particle fname aux adv prep museum La Joconde est actuellement au Louvre					

Table 10: Different versions of the answer A2663 (with *fname* for *first name* and *product(art)* for *artistic production*)

there are 6.17 answers per question (whatever the interaction modality). It averages to 4.12 over the written modality and 2.12 on the speech one. The difference comes from the fact that less people wanted to do the oral experiment (we have 99 participants for the web interface and 53 for the phone one) and that we have more unusable calls for the speech modality (bad audio quality, user hangs up before the end of the call,...). As a consequence, 2.8% of the questions were not answered orally (493 instead of 507 in total). The total corpus contains more than 23,000 words[3] and the speech corpus is more than one hour long. We observe that the number of words is twice larger in the written corpus (17,976) than in the speech corpus (7,128). Even if questions are the same on both modality, there is no ceiling effect. Words are counted from the raw data. The written corpus contains typos, misspellings, and abbreviations that make the word count bigger.

A detailed analysis of the average number of words per answer and duration of answer is presented in section 6.

For each answer, the modality and the type of the question are known. For each answer, a set of annotations is available. The next section details the answers annotations.

5. Corpus annotations and transformations

Several annotations and post-treatments were done on the corpus. We present them, showing the possible analyses they allow.

Observing the lemma Using the Tree-tagger (Schmid, 1994), we lemmatised the corpus (see table 10 line *Lemmatised*).

With such a version of the corpus, it is possible to observe the lexicon of the corpus and to compare the lexica depending on question features or interaction modality. For instance, it allows a comparison of speech and written lexica. (Garcia-Fernandez et al., 2009) shows that the lexicon is bigger for the written modality than for the speech one and that the word frequency is higher for the written modality than for the speech one. Moreover, observing the common lexicon of the two modalities, we show that common words are mainly function words, auxiliaries and modal verbs. We could conclude that the speech and written modalities use different vocabularies.

Moreover, comparing lexica depending on the semantic type of the question (quantity, location or time), (Garcia-Fernandez et al., 2009) shows that the lexicon is bigger for the quantity questions and highlights that estimations are less compact for quantity questions than for the others.

Observing the part-of-speech distributions A part of speech (POS) tagging was done using the Tree-tagger (Schmid, 1994). We substitute each word by its POS tag (see example table 10 line *POS*). This transformation makes it possible to observe the composition of the answers in terms of POS and more precisely to oppose function and content words. (Garcia-Fernandez et al., 2009) shows that spoken answers use proportionally more content words than written answers, so that spoken answers seem more focused on giving an information, while written answers are using more conjunctions and consist of more elaborated sentences.

[3] Here, a word is defined as a sequence of characters between spaces.

Question	Example
Q212	<focus>La Joconde</focus> <verb>se trouve</verb> <infoA>au Louvre</infoA> ? *<verb>Is</verb> <focus>the Mona Lisa</focus> <infoA>in the Louvre Museum</infoA>?*
Q258	Où <verb>est</verb> <focus>la Joconde</focus> ? *Where <verb>is</verb> <focus>the Mona Lisa</focus>?*

Table 11: Question annotation (with *infoA for information-answer*)

Answer	Example
A2879	<focus-pronoun>Elle</focus-pronoun> <verb> doit être </verb> au Louvre. *<focus-pronoun> It </focus-pronoun> <verb> should be </verb> in the Louvre.*
A155	Au <type>Musée</type> du Louvre, Paris. *In the Louvre <type>Museum</type>, in Paris.*
A2280	<focus>Une bouteille d'eau</focus> contient du liquide. (...) Si <focus-modified>la bouteille</focus-modified> contient 1 litre, <focus-pronoun>elle</focus-pronoun> <verb> pèsera </verb> un kilo et ainsi de suite. *<focus>A bottle of water</focus> contains liquid. (...) If <focus-modified>the bottle</focus-modified> contains 1 liter, <focus-pronoun>it</focus-pronoun> <verb> weights</verb> one kilo and so on.*

Table 12: Annotation of reuse from question in answer

Answer	Example
A2849	<iAnswer>Je ne suis pas sur</iAnswer>, il faut chercher dans un dictionnaire. [sic] *<iAnswer>I am not sure</iAnswer>, you should look in a dictionary.*
A155	<iAnswer>Au Muse du Louvre</iAnswer>, <iAnswer> Paris</iAnswer>. *<iAnswer>In the Louvre Museum</iAnswer>, <iAnswer>in Paris</iAnswer>.*

Table 13: Examples of information-answer annotation (with *iAnswer* for information-answer)

Observing the syntactic form Syntactic relation detection was produced using XIP, the Xerox Incremental Parser (Ait-Mokhtar et al., 2002). With these annotations (see table 10 line *Syntactic Parsing*), an analysis of the answer structure can be done. For instance, detecting recurrent syntactic structures gives information on different answer syntactic patterns which could be used for the surface generation in a QA system.

Observing the syntactico-semantic structure A multi-level automatic annotation of the corpus was also done providing information on extended named entities, question markers, and linguistic chunks (Rosset et al., 2007). This analysis is adapted to the question-answering task and is a non-contextual analysis (NCA). It gives information on the semantic structure of the answers. In the example table 10 line *NCA*, we observe that "Louvre" is recognised as a museum so we can check if this named entity type matches the one expected by the question. The same checking can be done regarding the verb: is the verb used in the answer a specific verb (verbs of location for instance, see section 3.5.), an auxiliary or an other type of verb? Moreover, this analysis makes it possible to detect dialogue acts such as expressions of misunderstanding (for instance "I didn't understand") which can help in distinguishing positive valence answers (answers which give an information answering the question) from negative valence answers (answers which do not contain any information answering the question).

Following sections describe manual annotations of the whole corpus.

Observing words from the questions being reused in the answer An annotation of the question elements which could be reused in the answer was done. Table 11 shows examples of annotation. For each question, we know its *focus*, its *principal verb*, the expected *type* of answer if explicitly named in the question (see for instance the three last examples of table 7), *additional information* to specify better the focus of the question, and the *information-answer* to be evaluated in the case of Yes-No questions (see *Yes-No question* in table 11).
An annotation of those elements in the answers was also done (see table 12). Three cases were considered concerning the *focus*: exact reuse, reuse with modification and pronominal reuse. Reuse with case modification, typos, abbreviations, and gender/number modifications are considered as exact reuses. Reuse of part of the focus are considered as reuse with modification. Synonyms are not considered as reuses. As we can see in the example A2280 of table 12, the focus can be reused in different ways in the same answer. We annotated a reuse of the *verb* whatever its realisation (tense, person, with a modal verb,...). Concerning the *type*, the different forms of units are considered equivalent ("cm", "centimeter", etc.).

Observing the element which answers the question We defined the information-answer as the shortest part of the answer which consists either (1) of a new information which corresponds to the question expected *general* type (in the table 13, "Paris" is an information-answer even if the precise type is "museum"), or (2) of an admission of

Type of additional element	Example
Irrelevance	**vas dans ta chambre :P** [sic] ***Go to your room :P***
Suggestion	Je ne suis pas sur, **il faut chercher dans un dictionnaire**. *I am not sure, **you should look in a dictionary**.*
Completion	Le 11 novembre 1918 **Rethondes** *November 11th 1918 **in Rethondes***

Table 14: Examples of answers containing aditionnal elements (in bold)

	All	Speech	Written	Open questions	Yes-No questions
Answers which reuse the focus...	22.54%	22.31%	22.65%	24.95%	17.76%
Answers which contain at least one exact reuse	62.48%	67.24%	60.16%	66.28%	51.38%
Answers which contain at least one reuse with modification	16.11%	14.41%	16.94%	16.66%	14.36%
Answers which reuse the focus only as a pronoun	23.39%	18.77%	25.63%	19.15%	35.91%

Table 15: Reuse of the question focus in the answers

incompetence (see table 13).
The information-answer is a key element in the answer and its annotation is useful for instance to observe its type, the number of information-answers in an answer and the relation between these information-answers.

Observing the additional elements Certain answers contain completions, suggestions or irrelevant elements. A manual annotation of these elements was done. Table 14 shows examples. A completion is defined as an element that gives additional information in relation with the question or the answer itself. A suggestion is defined as the expression of another way to find the information answering the question. Irrelevances are additional elements which are neither completions nor suggestions.
The annotation of additional elements makes it possible to remove them. Hence, an observation of the reduced answer is possible. But it also makes possible to observe additional elements more specifically, which could be useful for cooperative dialogue or question-answering systems.

6. Corpus analyses

In this section, we detail two analyses carried out on the corpus. The first one does not require any annotation or post-treatment of the corpus. It only takes into account available data on the duration and the size in word of the answers. The second analysis exploits the annotation of words being reused from the question, showing how the focus of a question is reused in the answers.

Duration and size of answers An analysis based on answer duration and size in words was conducted to characterize differences between speech and written modalities. The speech duration was measured by the speech detection system. For the written modality, duration was measured from the web page loading until the user clicked on “Validate the answer”. Answer size in words is calculated (see table 9) from the Tree-tagger results.
As a general observation we can say that subjects took in average more time to produce answers in writing (33 sec) than in talking (4.2). Written answers are in average longer (8.4 words) than speech ones (6). We can explain the difference in duration by the fact that on the written modality, our measure includes the time for reading the question and typing the answer whereas, on the speech modality, it starts when the subject starts speaking.

Statistical significance tests (two-sample Kolmogorov-Smirnov tests using the size or the duration as factor and modality as nominal) were carried out to measure the difference between *the distribution* of duration on speech and written modalities. We used the same test regarding the size of answers. They show that neither sizes ($p<0.0004$), nor duration ($p<0.00001$) of speech and written answers have the same distribution. Differences of distribution could be explained by the fact that subjects could be more or less familiar with keyboard, typing more or less quickly.
Differences in size show that humans produce longer answers while writing than speaking.

Which reuse of the question focus in the answer? Table 15 gives percentages of reuse of the question focus in the answers. Results are presented for the entire corpus and depending on the modality (speech *vs* written) and the type of question (open *vs* yes-no).

23% of the answers contain the question focus, whatever the kind of reuse (exact, with modification or with a pronoun). Among those answers, 63% contain the exact focus while 19% only refer to the focus using a pronoun.
Studying the corpus, we observe two kinds of focus reuse with modification. The first kind consists in reducing the phrase containing the focus to its head: “bouteille de lait” (“milk bottle”) is reused as “bouteille” (“bottle”). The second type consists in reducing the phrase containing the focus to the most semantically important word : “le mois de février” (“the month of February”) is reused as “février” (“February”).
The focus is more often replaced by a pronoun on the speech than on the written modality. It is also the case in answers to Yes-No questions compared to open questions.

7. Conclusion

We presented a corpus of natural language human answers and the way we acquired it.[4] Answers correspond to questions with fixed linguistic form, focus, and topic. Answers to a given question exist for two modalities of interaction: speech and written. The whole corpus of answers was annotated on different levels which allowed analyses from different points of view. A description of those analyses and annotations was presented. Two examples of corpus analyses are detailed. The first analysis shows some differences between speech and written modality especially in terms of length of the answers. The second analysis concerns the reuse of the question focus in the answers.

The corpus of questions is limited to 3 semantic types but the corpus may be extended to other question types. The questions were manually built but the protocol could be used with authentic questions (extracted from collaborative question-answering websites for example).

The analysis of this corpus will allow us to implement a set of rules to enhance the generation of answers in our question-answering system, both in the speech and written modalities.

8. Acknowledgements

This work has been partially financed by OSEO under the QUAERO program. We thank warmly Delphine Bernhard and Marie Guégan for them useful reviewing.

9. References

Salah Ait-Mokhtar, Jean-Pierre Chanod, and Claude Roux. 2002. Robustness beyond shallowness: incremental deep parsing. *Natural Language Engineering*, 8(2-3):121–144.

Olivier Ferret, Brigitte Grau, Martine Hurault-Plantet, Gabriel Illouz, Christian Jacquemin, Laura Monceaux, Isabelle Robba, and Anne Vilnat. 2002. How NLP can improve question answering. *Knowledge Organization*, 29(3-4):135–155.

Anne Garcia-Fernandez, Sophie Rosset, and Anne Vilnat. 2009. Collecte et analyses de réponses naturelles pour les systèmes de questions-réponses. In *Actes de TALN 2009*.

Daniel Luzzati. 2006. Essai de description interactive : l'exemple des questions quantificatrices. *Colloque La quantification*, 1:15.

David D. McDonald. 2003. Producing dialog at MERL: problems in generation engineering. In AAAI Spring, editor, *Proceedings of Natural Language Generation in Spoken and Written Dialogue*, pages 104–111.

Sophie Rosset, Olivier Galibert, Gilles Adda, and Éric Bilinski. 2007. The LIMSI participation to the QAst track. In Alessandro Nardi and Carol Peters, editors, *Working Notes of CLEF Workshop, ECDL conference*, Budapest, Hungary, September. Springer.

Helmut Schmid. 1994. Probabilistic part-of-speech tagging using decision trees. In *Proceedings of International Conference on New Methods in Language Processing*, volume 12. Manchester, UK.

Dave Toney, Sophie Rosset, Aurélien Max, Olivier Galibert, and Eric Bilinski. 2008. An evaluation of spoken and textual interaction in the RITEL interactive question answering system. In European Language Resources Association (ELRA), editor, *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, May.

TREC ciQA task. 2007. The TREC complex, interactive QA task.

[4] The corpus is freely available upon request