Assigning Wh-Questions to Verbal Arguments: Annotation Tools Evaluation and Corpus Building

Magali Sanches Duran [1], Marcelo Adriano Amâncio [1], Sandra Maria Aluísio [1]

[1] Center of Computational Linguistics (NILC) – Department of Computer Science
University of São Paulo – São Carlos-SP Brazil
magali.duran@uol.com.br, marcelousp@gmail.com, sandra@icmc.usp.br

Abstract

This work reports the evaluation and selection of annotation tools to assign wh-question labels to verbal arguments in a sentence. Wh-question assignment discussed herein is a kind of semantic annotation which involves two tasks: making delimitation of verbs and arguments, and linking verbs to its arguments by question labels. As it is a new type of semantic annotation, there is no report about requirements an annotation tool should have to face it. For this reason, we decided to select the most appropriated tool in two phases. In the first phase, we executed the task with an annotation tool we have used before in another task. Such phase helped us to test the task and enabled us to know which features were or not desirable in an annotation tool for our purpose. In the second phase, guided by such requirements, we evaluated several tools and selected a tool for the real task. After corpus annotation conclusion, we report some of the annotation results and some comments on the improvements there should be made in an annotation tool to better support such kind of annotation task.

1. Introduction

This paper reports the selection of an annotation tool to a specific annotation task: assigning wh-questions that link verbs to their arguments. For this, it is necessary to identify:

(a) verbs which give rise to questions;

(b) arguments which answer such questions and

(c) question labels that link properly (a) to (b).

For example, in *John died yesterday.* "Who?" is the question label that links the verb "died" to the argument "John". Similarly, "When" is the question label that links the verb "died" to the argument "yesterday". We use the term "argument" here in the same way it is used in the Propbank project (Palmer *et al.*, 2005), that is, on referring to both verbal arguments and modifiers.

There is a commercial system that annotates actions and arguments with wh-questions to support text mining[1]. However, the formalization of this task as it is reported here is new, to the best of our knowledge.

The corpus annotation task focused in this paper has been demanded by the project PorSimples (Aluísio *et al.*, 2008). In Section 2, we describe PorSimples trying to explain its interest in a corpus annotated with wh-questions. Readers interested in further details about PorSimples may access the project wiki page[2]. Section 3 briefly discusses some theoretical background. Section 4 presents the selected corpus, details the new task and reports the pilot test. Section 5 presents both a list of requirements for this new task and the evaluation of annotation tools which had preceded the annotation task. Section 6 analyses the performance of the selected tool. Future work is addressed in Section 7.

2. PorSimples: Adapting Web content for low-literacy readers.

The main goal of PorSimples is to develop natural language processing (NLP) technologies related to Text Adaptation (TA) in order to promote digital inclusion and accessibility for people with low levels of literacy.

Text adaptation is a very well known practice used in educational settings. Young (1999) mentions two different techniques for text adaptation: Text Simplification and Text Elaboration. The first can be defined as any task that reduces the lexical or syntactic complexity of a text while trying to preserve meaning and information. Text Elaboration aims at clarifying and explaining information and making connections explicit in a text, using, for example, definitions, synonyms or hypernyms of the text words.

Since 2001, the INAF index (National Indicator of Functional Literacy) has been annually computed to measure the levels of functional illiteracy of Brazilian population. The 2009 report presented a still worrying scenario: 7% of individuals were classified as illiterate; 21% as literate at rudimentary level; 47% as literate at basic level; and only 25% as literate at advanced level (Montenegro, 2009). Thus, especially when we consider that a large portion of Brazilian people (about 28%) is functional illiterate, we argue that an assistive technology for adapting web content is an urgent necessity for digital inclusion of low literacy people.

To overcome this scenery PorSimples is developing a text adaptation system to allow poor literacy readers and children or adults in literacy process to understand Web content and to develop comprehension and critical reading skills. PorSimples uses methods to support two styles of reading:

- skimming for getting the gist of a Web text, via summarization and text simplification methods in FACILITA system (Watanabe et al., 2009) and

---

[1] http://www.cortex-intelligence.com/tech/
[2] http://caravelas.icmc.usp.br/wiki/index.php/Principal

- detailed reading, by using text presentation schemes, which should highlight the named entities of a text and perform lexical elaboration, in FACILITA EDUCATIVO system (Watanabe et al., 2010).

In this paper we present a new method of detailed reading by exhibiting questions to make clear semantic relations that link verbs to their arguments.

PorSimples makes the assumption that exhibition of questions which link verbs to their arguments will help users to develop a strategy for reading comprehension, although a solution for cognitive problems obviously encloses more complex measures.

The annotated corpus focused herein will be used as a training corpus by PorSimples to train a classifier for automatic assignment of questions.

3. Theoretical Background

The assignment of question labels linking verbs to their arguments is an issue that has much in common with question-answering systems (Q&A) (for Portuguese, see Bick, 2003). Notwithstanding, in our context: a) answers are not unknown, as they are within the sentence being read; b) questions are not made by users and c) the purpose of the questions is not to find answers, but to help reading comprehension.

Literature usually points out contribution of semantic role labelling (SRL) to Q&A challenges. When we analyzed two SRL projects for English – Framenet (Baker et al. 1998) and Propbank (Palmer et al. 2005), we realized a strong correlation between question labels and semantic role labels, especially those related to adjunct roles as *time*, *locative*, *manner*, *purpose*, *cause*, *direction* and *quantity*. Reports on SRL have been used to identify annotation tools likely to meet our target.

We realized the possibility to take profit of semantic role labeling (SRL) to map the question labels. As SRL is largely discussed in NLP community, it would be logical to start by SRL and deriving question labeling as automatically as possible. Notwithstanding, there is no corpus in Portuguese annotated with SRL. Implementing such kind of annotation in a Treebank of Portuguese would be very time-consuming and would affect the schedule of PorSimples. For this reason, we decided to start from wh-questions annotation, paying attention to details that make it possible to take profit of such annotation in future works related to SRL annotation.

4. Corpus and Annotation Task

Question annotation was performed under a corpus of simplified texts (Caseli et. al, 2009) downloaded from Portal of Parallel Corpora of Simplified Corpus[3] .

There are two main reasons considered here to annotate a simplified corpus:

- Simplified texts consist of active sentences, have no relative clauses, no appositions and have few coordinate and subordinate clauses, features which

made them less exposed to automatic parsing errors. This is intended to ensure a better performance in the automatic steps of pre-annotation process as well as to provide a better input for the future steps of learning rules;

- Simplification rules used to generate the texts of the corpus (Specia *et al.*, 2008) did not produce changes relating to adjuncts, that is, they do not include losses of relevant material for the intended annotation.

This corpus has been previously annotated by the parser Palavras (Bick, 2000), but syntactic annotation has not been submitted to human correction.

We describe some detail about this corpus in the next sub-section.

4.1 Selected Corpus

Six corpora covering two different genres and three levels of literacy were compiled as part of the PorSimples project. Texts were manually simplified by a linguist, expert in text simplification, according to the two levels of simplification: natural and strong. The first type results in texts adequate for people with basic literacy level and the second type results in texts adequate for people with rudimentary level. The difference between these two is the degree of application of simplification operations to complex sentences. For strong simplification, operations are applied to all complex syntactic phenomena present in the text in order to make it as simple as possible, while for natural simplification these operations are applied selectively, only when the resulting text remains "natural".

The first corpus is composed of 104 general news articles from Brazilian newspaper Zero Hora (ZH original). The other corpus is composed of popular science articles from Caderno Ciência (CC original), which is a section of Folha de São Paulo, a mainstream Brazilian newspaper.

We decided to start with ZH corpus although the annotation of CC corpus is already underway. Moreover, we have chosen the strong simplified version of simplification. Table 1 shows a few statistics about these six corpora.

Corpus	Doc	Sent	Words	Avg. words per text (std. deviation)	Avg. words p. sentence
ZH original	104	2184	46190	444.1 (133.7)	21.1
ZH natural	104	3234	47296	454.7 (134.2)	14.6
ZH strong	**104**	**3668**	**47938**	**460.9 (137.5)**	**13.0**
CC original	50	882	20263	405.2 (175.6)	22.9
CC natural	50	975	19603	392.0 (176.0)	20.1
CC strong	50	1454	20518	410.3 (169.6)	14.1

Table 1: Corpus statistics.

4.2 Defining the Annotation Task

We elaborated question annotation guidelines and defined previously 44 question labels, exceeding the number of interrogative pronouns in Portuguese – "quem", "que"/"qual", "quando", "quanto", "onde" and "como" (who, what, when, how much, where and how).

3 http://caravelas.icmc.usp.br/portal/index.php

This is because in Portuguese prepositions that follow verbs or introduce adverbial clauses are moved to the beginning of interrogative clauses, motivating prepositioned interrogative labels as "com quem" (with+who), "de onde" (from+where), "por quanto" (for+how much). Except for "como", all Portuguese interrogative pronouns may be combined with prepositions. For example, an affirmative clause like "Ele se preocupa com seus filhos" (He worries about his children) may give rise to an interrogative clause like "Com quem ele se preocupa?" (Who does he worry about?). Compound question labels include "por quê" (why), which asks for causes and "para quê" (what for), which asks for purposes. We also included "quanto tempo" (how long) and "com que frequência" (how often) based on Hagège *et al*. (2007), who points out three types of temporal expressions: calendar time (when), duration (how long) and frequency (how often).

4.3 Pilot Test

As this is a new task, we decided to make a pilot test using an annotation tool already mastered by members of our group. The pilot test aimed at:

- measuring feasibility and reproducibility of the task;
- writing evidence-based guidelines to support the selection of an appropriate annotation tool.

We used MMAX2[4] (Müller and Strube, 2006) because some members of our group had previous experience with it. In this way, we shortened time required to learn how to use and customize the tool.

Our choice was also motivated by the multi-level feature of MMAX2, which allowed us to create a level for sentence segmentation, a level for verb segmentation and a level for answer (argument) segmentation. Question labels were configured as attributes of answer's segments. This enabled us to organize annotation process in several steps, reducing errors occurrence, detection and correction.

In this phase, segmentation of answers (that is, arguments) has been performed automatically by selecting principal nodes of syntactic trees. Such task was followed by human revision.

Afterwards, seven annotators performed the assignment of questions labels. Kappa inter-annotator agreement was of 0,78, which is a good result considering annotators were not specialists in Linguistics and had only 15 minutes to read guidelines and 15 minutes to ask questions before starting the task.

Pilot test results showed us the task is feasible and reproducible, but evidenced some problems that may rise if annotation tool is not suitable for the task.

We found some drawbacks in MMAX2 for our purpose:

- There is no support for including or deleting markables after project creation, that is, markables

and attributes cannot be changed during annotation. This feature limits the flexibility of annotation process;

- There is no support for creating schemas: it is required an XML expert to configure the tool before human annotation task;

- It is not possible to assign more than one label to each segment on a same level. This is an important requirement for us, as in some cases a segment admits two different questions;

- There is no UNDO option: to undo an action requires to do the previous action again;

- There is no means to restrict access to segmentation while performing a categorization task, although it is possible inactivating a markable not being currently in use (but this is different from restricting access); such an option would avoid segment creation or deletion by an involuntary click;

- Annotation window does not show which segments have already been labelled.

5.   Annotation Tools Evaluation

The pilot test allowed us to observe desirable and undesirable features of an annotation tool for our task. To guide our choice, we elaborated the following list of requirements:

- labels and attributes edition during annotation task

- multi-level annotation;

- multi-level search engine;

- annotation on parse trees;

- comments edition during annotation;

- whole visualization of segments already labelled;

- configuration of user's rights to read and edit labels;

- sub-specification of labels;

- graphical interface;

- easy label selection.

While searching for annotation tools that satisfy these requirements, we find Tred[5], Palinka[6], Knowtator[7], UAM[8] and SALTO[9]. We analyzed their features, following our list of requirements and observing as well whether they are or not free available and recently updated. The result is presented on Table 2.

As may be observed, none of them satisfies all of our requirements. For this reason, we decided to discard firstly the tools not being updated and, then, from the other tools, we selected two that suit our needs best: SALTO and UAM.

---

[4] mmax2.sourceforge.net/

[5] http://ufal.mff.cuni.cz/~pajas/tred/

[6] http://clg.wlv.ac.uk/projects/PALinkA/

[7] knowtator.sourceforge.net

[8] http://www.wagsoft.com/CorpusTool/

[9] http://www.coli.uni-saarland.de/projects/salsa/ page.php?id=software

	MMAX 2	NITE	SALTO	UAM	PALINKA	KNOWTAT	TRED
free	x	x	*	x	x	x	x
recent upgrades	x	x	x	x	-	-	x
annotation on parse trees	-	-	x	-	-	-	x
labels and attributes edition during annotation task	-	-	x	x	-	x	x
graphical interface	x	x	x	x	x	x	x
whole visualization of segments already labelled	-	x	x	x	x	x	x
easy label selection	x	x	x	x	x	x	x
sub-specification of labels	x	x	-	x	-	x	x
configuration of user's rights	-	x	x	-	-	-	-
comments edition during annotation	-	-	x	x	-	x	-
multi-level search engine	x	x	-	x	-	x	x
multi-level annotation	x	x	-	x	-	x	x

"x" = satisfies the requirement

"-" = does not satisfies the requirement

* SALTO is free for research purposes upon request.

Table 2 – Comparison of Annotation Tools.

SALTO (Burchardt et al. 2006) was developed for SRL annotation in a corpus of German, following FrameNet's methodology. It allows annotation directly on parse trees, is friendly and easy to configure. From features listed as desirable, the only ones SALTO does not satisfy are: multi-level annotation and user-friendly searching engine which supports parameters of several levels simultaneously (TIGERSearch is incorporated into SALTO and supports queries to the corpus, but is not user-friendly).

UAM annotation tool is more generic purpose when compared to SALTO; it is multi-level and supports searches as we required, besides allowing us to label whole texts as well as parts of texts. However, UAM does not allow annotation on parse trees, feature that would facilitate our task, considering syntactic parsing information is essential to define segments which will receive question labels. UAM presents an interface user-friendly and allows changing annotation schema during annotation task (this was not possible in our previous experience with MMAX2). We analyzed both tools, SALTO and UAM and decided to adopt SALTO due to its graphical annotation mode, visual editor, mouse-menus and annotation on parse trees, features which made our annotation task easier, faster and more comfortable.

Besides that, SALTO has two types of user modes: user(s) (who does the annotation) and administrator(s) (who supplies corpora and controls annotation results). This feature was not listed as a requirement for our task, but may be useful in future works.

When we had already initiated the corpus annotation, we received a suggestion about a more comprehensive annotation tool, NITE[10] Toolkit. We analyzed it in order not to loose the opportunity of finding a more appropriate tool for our task. NITE is a multimodal corpus annotation tool that meets several of our requirements: it is multi-level, open source and updated. More important than this, its query language allows information from different media annotation to be treated as one coherent set. Such resource makes it possible, for example, to combine text annotation with video annotation, an unimaginable but necessary resource for those who analyze speech and gesture synchronic relations.

In our task, however, it is important to see the syntactic tree during annotation task, not simply to see syntactic and semantic annotations combined through the query language after annotation process. For this reason, we kept our decision on using SALTO annotation tool.

6. Evaluation of the Selected Tool and Task

In order to evaluate the efficiency of our selection process and to register improvements required to better adjust an annotation tool to wh-question annotation we present some comments about the performance of SALTO during the task.

From 3668 sentences annotated with SALTO (corpus ZH strong), 182 sentences (5%) were disregarded because of grave parsing errors (we flagged them as "Wrong Subcorpus"). We flagged also 356 sentences that presented minor parsing errors: we will try machine learning with and without them. Such decision intended to select only sentences with correct parsing analysis as input for statistical machine learning. In this way, we made some quality control on sentences to compensate for the lack of human revision on parsing analysis.

We assigned 10.438 question labels, distributed as follows:

"o quê?-DIR"	2862	27,42%
"quem?-ESQ"	2211	21,18%
"o quê?-ESQ"	1233	11,81%
"onde?"	864	8,28%
"quando?"	758	7,26%
"como?"	244	2,34%
"quem?-DIR"	189	1,81%
"qual?-ESQ"	162	1,55%
"para quê?"	159	1,52%
"como?-verbal"	143	1,37%
"de quê?"	127	1,22%
"em quê?"	109	1,04%
"a quê?"	100	0,96%
"por quê?"	80	0,77%
"a quem?"	76	0,73%
"que idade?"	68	0,65%
"quanto?"	63	0,60%
"de onde?"	64	0,61%
"em que condicão?"	59	0,57%

---

[10] http://groups.inf.ed.ac.uk/nxt/

"para quem?"	56	0,54%
"com o quê?"	62	0,59%
"com quem?"	52	0,50%
"aonde?"	51	0,49%
"com que frequência?"	50	0,48%
"de quanto?"	46	0,44%
"até quando?"	38	0,36%
"ha quanto tempo"	33	0,32%
"para onde?"	33	0,32%
"para o quê?"	32	0,31%
"por quem?"	35	0,34%
"pelo que?"	29	0,28%
"quanto tempo?"	28	0,27%
"de quem?"	26	0,25%
"por onde?"	28	0,27%
"desde quando?"	22	0,21%
"quais?-ESQ"	21	0,20%
"de onde?-filiação"	20	0,19%
"a quanto?"	19	0,18%
"sobre o quê?"	16	0,15%
"quantos?"	15	0,14%
"por quanto tempo?"	14	0,13%
"como o quê?"	14	0,13%
"para quanto?"	14	0,13%
"em quanto?"	12	0,11%
"em quanto tempo?"	11	0,11%
"até onde?"	10	0,10%
"contra quem?"	8	0,08%
"em quem?"	7	0,07%
"em que período?"	7	0,07%
"em que direcão?"	8	0,08%
"com que consequência?"	6	0,06%
"para quando?"	6	0,06%
"contra o quê?"	4	0,04%
"por quanto?"	4	0,04%
"até quanto?"	4	0,04%
"depois de quanto tempo?"	4	0,04%
"em que língua?"	4	0,04%
"a partir de onde?"	4	0,04%
"de quando?"	2	0,02%
"que lugar?"	2	0,02%
"a partir de quando?"	2	0,02%
"por que distância?"	2	0,02%
"para com quem?"	2	0,02%
"entre o quê?"	2	0,02%
"com quanto?"	2	0,02%
Total labels assigned:	10438	100,00%

Table 3: Assigned Labels.

As may be observed in Table 3, there are 65 question labels. When configuring these labels, we tried to foresee their mapping to role labels. For example, we created two labels "quem": "Quem?-DIR", related to Arg1 or Arg2 of Propbank role labels (syntactic role: direct object), and Quem?-ESQ" related to Propbank's Arg0 or Arg1 (syntactic role: subject). Except for role labels associated to subject and direct object, question labels present greater granularity than Propbank role labels. For the semantic role of place, for example, there are nine question labels: "onde?", "de onde?", "aonde?", "para onde?", "por onde?", "de onde?-filiação", "até onde?", "a partir de onde?", "que lugar?".

Wh-questions have been configured as frame elements to facilitate our task. Most frequent wh-questions are exhibited automatically when we evoke our unique "frameset": "pergunta"; other questions rest hidden and may be selected by clicking on right button (Figure 1).
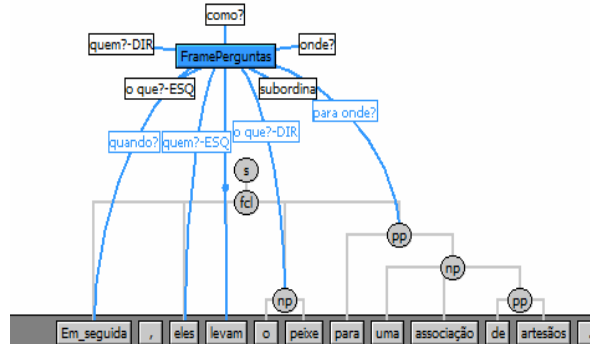

Figure 1: Annotation Screen.

Many times, the verb that "evokes" the questions is not constituted of a single verb. There are cases in which the occurrence of auxiliary verbs composes Verbal Phrases (VP) with 2, 3, 4 and even 5 verbs in sequence. Such cases show over-auxiliarity phenomena, that is, more than one type of auxiliary verb occurs in a same sequence. For example:

> zh003.s34: Fechar as pontes , e nada mais , **vai acabar colocando** os moradores de rua nas praças e na frente dos prédios .
>
> Question: **O que** *vai acabar colocando*?
> Answer: Fechar as pontes e nada mais.
> Question: *Vai acabar colocando* **quem**?
> Answer: Os moradores de rua.
> Question: *Vai acabar colocando* **onde**?
> Answer: Nas praças e na frente dos prédios.

As may be seen in this example, we annotated "vai acabar colocando" as a whole verbal phrase (VP) and, for this reason, the VP takes part of questions made. In Portuguese, "vai" is a form of verb "ir" and this verb, when followed by an infinitive verb, conveys a future sense, like "going to" in English. Relating to "acabar", when followed by a gerund verb, encodes a resultative aspectual sense, that is, the notion of culmination of an event.

The phenomena of over-auxiliarity led us to revise the concepts of temporal, modal, aspectual and passive voice auxiliarity in order to interpret data found in corpus and make decisions about how to manage them.

In the same way, sometimes the verb that "evokes" the questions is constituted of a light verb plus a noun (phenomena known in Portuguese as support verbs constructions). For this, it is necessary to decide, during the annotation task, which elements should be marked as a whole. This is very important to our purpose, as every word that pertains to the evocating node will take part in the questions.

The next example shows the verb "dar a entender", which means "to suggest".

---

zh024.s12: Todas_as regras do desfile **dão a entender** que a comissão organizadora quer um desfile mais dócil .

Question: **Quem** *dá a entender*?
Answer: Todas as regras do desfile
Question: *Dá a entender* **o quê**?
Answer: Que a comissão organizadora quer um desfile mais dócil.

Figure 2 shows how we "evoked" "Frame Perguntas" joining these three tokens "dão a entender".
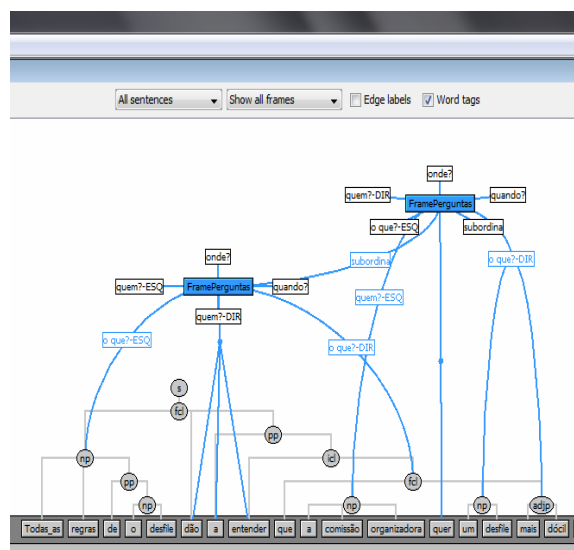

Figure 2: Annotation of a multiword verb.

As shown in Table 4, we annotated 4932 verbs and assigned 10438 question labels, an average of 2,12 labels by verb.

Analysed sentences	3486
Annotated verbs	4932
Simple verbs	4159
Multiword verbs	773
Wh-question labels assigned	10438

Table 4: Statistics.

As we annotated a corpus of simplified texts, it was expected not to find sentences with many verbs. This has been confirmed (see Table 5), as 69% of the sentences presented only one verb, 21% two verbs and 6% three verbs. Lower complexity of simplified sentences made our task simpler and faster.

Sentences with 1 verb	2401
Sentences with 2 verbs	809
Sentences with 3 verbs	204
Sentences with 4 verbs	61
Sentences with 5 verbs	9
Sentences with 6 verbs	2
Total	3486

Table 5: Number of verbs by sentence.

In spite of using SALTO for a purpose different from that it was made for, we have been successful.

A feature that allows frames to be elements of another frames was very useful, although it had not been previously required. Such feature enabled us to create a frame element called "subordinada", which we used to annotate subordination relationships.

We did not used SALTO facilities to distribute the task among several annotators. In spite of that, if we had had more time and resources, we surely would have taken profit of such feature. The experience showed us it is a hard work to revise all question annotations. The better way is to compare automatically two or three annotators' decisions on the same task and focus revision on annotations which don't match.

As we had only one annotator, SALTO facility to create new labels during annotation task was very useful, making annotator independent of programmer's interventions.

A facility desirable in an annotation tool, but not implemented in SALTO, is to search and to substitute labels. Such operations need to be executed by using TIGER searching tool, which is not user friendly (for a linguist, at least).

In the same way, it would be very useful a facility to select sentences following given criteria, including keywords, syntactic features or question labels assigned.

For example: select all the sentences that present the question label "quando". Such facility would enable the linguist annotator to review annotations that share some features.

Annotation guidelines, previously elaborated, were enriched during annotation process, as they incorporated several decisions about unexpected occurrences.

There are many decisions to be made during the annotation task. For example, there are cases of two semantic arguments that are parsed as a unique syntactic segment. The contrary also occurs, and we have to decide when to split and when to join syntactic segments.

Some arguments have no suitable question label to be assigned. In these cases, we decide to create new labels during annotation task. In case of performing annotation task simultaneously by more than one annotator, new labels creation should be inhibited and a label "no label found" should be created to register annotators' doubts.

7. Future Work

We are currently annotating CC corpus and, until July 2010, the training corpus will be made publicly available at PorSimples site.

Regarding PorSimples, the next task is to use the annotated corpus as a training corpus for statistical machine learning, aiming at automatic assignment of wh-questions.

The corpus will also be used to map semantic roles in a new project aiming to build the PropBank.Br (a Proposition Bank for Brazilian Portuguese language). One may also take profit from this annotated corpus to improve question-answering systems.

8. Acknowledgements

9. References

Aluísio, S., Specia, L., Pardo, T., Maziero, E., Fortes, R. (2008). Towards Brazilian Portuguese Automatic Text Simplification Systems. In: *Proceedings of The Eight ACM Symposium on Document Engineering (DocEng 2008),* pp. 240-248.

Baker, C.F., Fillmore, C. J., Lowe, J.F. (1998). The Berkeley FrameNet Project. In: *Proceedings of Computational Linguistics 1998 Conference*, University of Montréal: Association for Computational Linguistics, pp. 86-90.

Bick, E. (2003). A Constraint Grammar Based Question Answering System for Portuguese. In: Fernando Moura Pires & Salvador (eds.) *Progress in Artificial Intelligence* (Proceedings of EPIA'2003, Beja, Dec. 2003), Springer, pp. 414-418.

Bick, E.: The Parsing System Palavras Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework. Aarhus, Denmark, Aarhus University Press. (2000).

Burchardt, K. Erk, A. Frank, A. Kowalski and S. Pado. (2006).SALTO - A Versatile Multi-Level Annotation Tool. In: *Proceedings of LREC-2006*, Genoa, Italy.

Caseli, H.M.; Pereira, T.F., Specia, L.; Pardo, T.A.S.; Gasperin, C.; Aluísio, S.M. (2009). Building a Brazilian Portuguese parallel corpus of original and simplified texts. In Alexander Gelbukh (ed), *Advances in Computational Linguistics, Research in Computer Science*, 10th Conference on Intelligent Text Processing and Computational Linguistics (CICLing-2009), March 01–07, Mexico City. v. 41, pp. 59-70.

Hagège, C., Baptista, J., Mamede, N. (2007). Proposta de Anotação e Normalização de Expressões Temporais da Categoria TEMPO para o HAREM II. http://www.linguateca.pt/aval_conjunta/HAREM/TEMPO_ 2008_02_18.pdf.

Montenegro I. P. and Educativa A. (2009) INAF Brasil - 2009 - Indicador de Alfabetismo Funcional, 2009. Available at: http://www.ibope.com.br/ipm/relatorios/

Müller, C., Strube, M. (2006). Multi-level Annotation of Linguistic Data with MMAX2. In: Braun, S., Kohn, K., Mukherjee J. (eds.) *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, Frankfurt a.M.: Peter Lang, pp. 197--214.

Palmer, M., Gildea, D., Kingsbury, P. (2005). The Proposition Bank: A Corpus Annotated with Semantic Roles, *Computational Linguistics Journal*, 31:1.

Specia, L., Aluisio, S.M., Pardo, T.A.S.: Manual de Simplificação Sintática para o Português. Technical Report NILC-TR-08-06, São Carlos-SP. (2008).

Watanabe W.M., Candido Jr. A, Uzêda V., Fortes R. P. M., Pardo T. A. S., Aluisio S. M. (2009) Facilita: reading assistance for low-literacy readers. In the Proceedings of ACM SIGDOC 2009 - ACM International Conference on Design of Communication, 2009, Bloomington, IN. v. 1. p. 29 - 36.

Watanabe, W. M.; Candido Jr. A.; Amancio, M. A.; Oliveira, M.; Pardo, T. A. S.; Fortes, R. P. M.; Aluísio, S. M. (2010) Adapting web content for low-literacy readers by using lexical elaboration and named entities labeling. p. 1-9. To be published in the 7th International Cross-Disciplinary Conference on Web Accessibility (http://www.w4a.info/), 26th, & 27th, April, 2010, Raleigh, NC, USA.

Young, D. N. (1999) Linguistic simplification of SL reading material: effective instructional practice? The Modern Language Journal, 83(3):350–366.