HAL
open science

Constraint-Based Parsing as an Efficient Solution: Results from the Parsing Evaluation Campaign EASy

Tristan Vanrullen, Philippe Blache, Jean-Marie Balfourier

To cite this version:

Tristan Vanrullen, Philippe Blache, Jean-Marie Balfourier. Constraint-Based Parsing as an Efficient Solution: Results from the Parsing Evaluation Campaign EASy. 2006, pp.165-170. hal-00135423

HAL Id: hal-00135423

https://hal.science/hal-00135423

Submitted on 7 Mar 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Constraint-Based Parsing as an Efficient Solution: Results from the Parsing Evaluation Campaign EASy

Tristan Vanrullen, Philippe Blache, Jean-Marie Balfourier

Laboratoire Parole et Langage / CNRS UMR 6057
23 Avenue Robert Schuman
13100 Aix-en-Provence France
{tristan.vanrullen, philippe.blache}@lpl.univ-aix.fr

Abstract

This paper describes the unfolding of the EASy evaluation campaign for french parsers as well as the techniques employed for the participation of laboratory LPL to this campaign. Three symbolic parsers based on a same resource and a same formalism (Property Grammars) are described and evaluated. The first results of this evaluation are analyzed and lead to the conclusion that symbolic parsing in a constraint-based formalism is efficient and robust.

1. Introduction

Recent advances in parsing technologies make it possible to deal with different kinds of inputs: several shallow parsing systems are now available as well as deep parsers, providing a set of solutions according to the needs. Parsing unrestricted texts will make use of shallow parsers whereas deep techniques can be used when fine-grained parses are needed. However, these techniques do not make it possible to reuse the same technology and the same resources both for deep and shallow parsing, and offering the possibility to choose the parsing granularity (see workshop deep & shallow parsing). Such technology has been proposed in the framework of Property Grammars (see (Blache, 2001 and 2005)). PG is a fully constraint-based formalism in which linguistic information is entirely specified in terms of non hierarchical constraints. A grammar in this approach is a set of constraints and parsing an input consists in evaluating this set. The input is then described (we say characterized) by the set of satisfied and violated constraints after such an evaluation. This conception of parsing makes it possible to treat any kind of input, whatever its form. In terms of parsing, we know that symbolic techniques offer several well-known advantages in terms of grammar development as well as reusability of the components. We show in this paper that these techniques can also be robust and efficient. In a former study (see (Vanrullen and Blache, 2002)), due to the lack of French treebanks at this time, we have proposed an evaluation technique based on the comparison of the parsers outputs without needing a reference corpus. That kind of evaluation lead us to improve our parsers by detecting their differences and by widening our knowledge of their weaknesses in terms of robustness, precision and efficiency. The French evaluation campaign *EASy* (Evaluation des Analyseurs Syntaxiques, see, (Vilnat & al, 2003 and 2004)) gave us the possibility to evaluate our parsers among several participants and with a strict evaluation framework based on a guideline and a large reference corpus.

This paper gives an account of the evaluation within the EASy campaign of three Property Grammar parsers, reusing the same mechanisms and the same resources (lexicon and grammar).

After a brief introduction to the EASy evaluation framework and a short presentation of Property Grammars, we will describe the three parsers based on Property Grammars build in *Laboratoire Parole et Langage* (hereafter LPL) to participate to the EASy campaign. Then, we will show our results and discuss them.

2. The evaluation framework

The EASY project (see http://www.elda.org/easy) aims at the evaluation of parsers for French. It proposes an evaluation methodology making it possible to compare syntactic analyzers and, as a side effect, produce a large validated linguistic resource by combining automatically the results of the campaign.

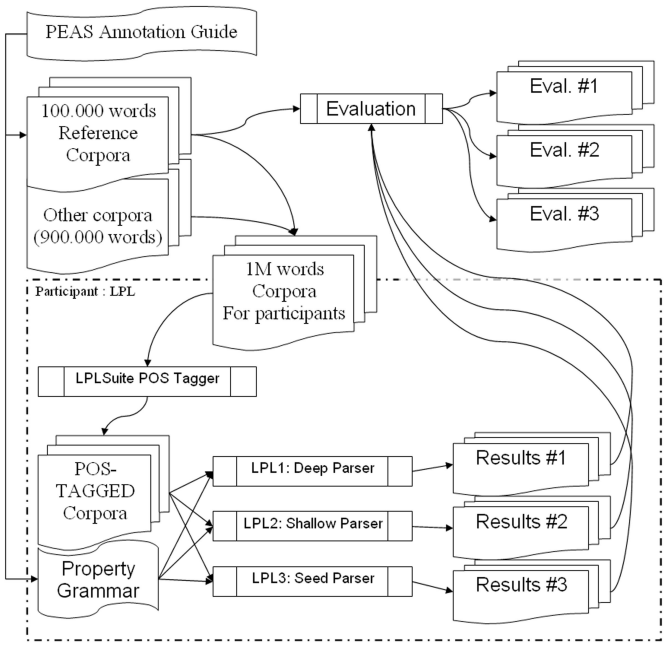

Figure 1: EASy campaign and three LPL parsers based on PGs

The corpus set (1 million words) comes from different sources newspapers, questions, websites, oral transcriptions, etc.), and contains different types: general corpora (21%), literary corpora: (23%), mail (15%), medical (6%),

oral transcriptions (28%) and questions (7%). Each kind of corpus contains a reference bracketed part. The project comes with an evaluation framework for many different approaches.

The main idea of this campaign relies on the use of a common target format for all the evaluated techniques. This format (PEAS, see (Gendner & al, 2002, 2003 and 2004)) is defined as a syntactic annotation standard. The EASy campaign is finishing, this paper reports the first results concerning bracketing.

The figure 1 shows the main steps of the EASy campaign and some elements of our participation. After the PEAS guide was defined, a part of the million words was manually bracketed (10%). This part -unknown to participants- has been used as reference for the evaluation process.

The whole corpus was transmitted to participants. A task of POS-tagging was needed to prepare parsing. Our three parsers were then used to analyze the tagged corpus and we transmitted three sets of results to the EASy organizers.

For our participation, we used our NLP Framework (LPL-suite) which allows basic tasks such as tokenizing, POS-tagging and lexical information retrieval. Our lexicon of 440.000 entries covered the main part of the corpus. Thus, few manual tasks were needed to prepare the parsing process.

Another preliminary task was to build our own Property Grammar based on the PEAS guide. For each specification given in this guide, as for the example hereunder extracted from the PEAS guide, we translated the whole guide into constraints according to the Property Grammars specification.

<u>Noun Phrases</u> *contain a noun which may be preceded by a determiner (...) and/or an adjective with its modifiers (which can contain adverbs), or a proper noun or a non clitic pronoun. When an adjective is preceded by a determiner, it stays an adjective but belongs to a noun phrase.(...) When several proper nouns follow each other without determiner nor preposition, they belong to the same noun phrase (...).*

A particularity of this campaign and the PEAS guide is that the syntactic structure of the reference corpus is not recursive nor hierarchical. All participants have to follow this guide in order to evaluate their outputs . We will see in the section 5. that some elements of the guide stayed ambiguous or difficult to translate. One of the interests of symbolic techniques can be found here due to the fact that it was possible for us to build a specific grammar instead of adapting our tools to the goal. This fact is very important in our approach because many different grammars can be developed without modifying the parsers.

A last particularity of the EASy evaluation framework required a development effort: the evaluated parsers outputs had to be deterministic. Only one syntactic analysis per sentence is compared to a reference bracketed sentence. This fact implied the development of determinization tools for our two non-deterministic parsers (see section 4.).

3. Property Grammars

We present in this section the formalism of Property Grammars (see (Bès, 1999) for preliminary ideas, and (Blache, 2000 and 2005) for a precise presentation). The main characteristics of Property Grammars (noted hereafter PG), is that all information is represented by means of constraints. Moreover, grammaticality does not constitute the core question but becomes a side effect of a more general notion called characterization: an input is not associated to a syntactic structure, but described with its syntactic properties.

PG makes it possible to represent syntactic information in a decentralized way and at different levels. Instead of using sub-trees as with classical generative approaches, PG specifies directly constraints on features, categories or set of categories, independently of the structure to which they are supposed to belong. This characteristic is fundamental in dealing with partial, underspecified or non canonical data. It is then possible to stipulate relations between two objects, independently from their position in the input or into a structure. The description of the syntactic properties of an input can then be done very precisely, including the case of non canonical or non grammatical input. We give in the remaining of the section a brief overview of PG characteristics

All syntactic information is represented in PG by means of constraints (also called properties). They stipulate different kinds of relation between categories such as linear precedence, imperative co-occurrence, dependency, repetition, etc. There is a limited number of types of properties. In the grammar we developped for the EASy campaign, we used the following ones:

Linear precedence $Det < N$ (a determiner precedes the noun)

Dependency $AP \rightarrow N$ (an adjectival phrase depends on the noun)

Requirement $V[inf] \Rightarrow to$ (an infinitive comes with *to*)

Exclusion $seems \neq ThatClause[subj]$ (the verb *seems* cannot have *That* clause subjects)

Uniqueness $Uniq_{NP}\{Det\}$(the determiner is unique in a *NP*)

Obligation $Oblig_{NP}\{N, Pro\}$(a pronoun or a noun is mandatory in a *NP*)

This list can be completed according to the needs or the language to be parsed. In this formalism, a category, whatever its level is described with a set of properties, all of them being at the same level and none having to be verified before another.

Parsing a sentence in PG consists in verifying for each category the set of corresponding properties in the grammar. More precisely, the idea consists in verifying for each constituent subset its relevant constraints (i.e. the one applying to the elements of the subset). Some of these properties can be satisfied, some other can be violated. The result of this evaluation, for a category, is a set of properties together with their evaluation. We call such set the characterization of the category. Such an approach makes it possible to describe any kind of input.

Such flexibility has however a cost: parsing in PG is theoretically exponential (see (Vanrullen, 2005)). This complexity comes from several sources. First, this approach offers the possibility to consider all categories, independently from its corresponding position in the input, as possible constituent for another category. This makes it possible for example to take into account long distance or non projective dependencies between two units. Moreover, parsing non canonical utterances relies on the possibility of building characterizations with satisfied and violated constraints. In terms of implementation, a property being a constraint, this means the necessity to propose a constraint relaxation technique. Constraint relaxation and discontinuity are the main complexity factors of the PG parsing problem.

4. The parsers

Three PG parsers (hereafter LPL1, LPL2 and LPL3) implementing different parsing strategies (from shallow to deep, from deterministic to non-deterministic, from flat to hierarchical) have been involved in the campaign. These parsers, even though significantly different in their conception, all rely on constraint satisfaction. Moreover, they all use the same resources: a lexicon and a PG French grammar which has been designed following the PEAS requirements.

LPL1 is a deep non-deterministic parser developed in the Delphi programming language. It produces all possible parses for a given input. It allows constraints to be relaxed in case of ill-formed input. Selecting the best parse (because of the need for a deterministic output in the EASy campaign) makes use of a determinization algorithm applied after the parsing process: the chosen characterization is the one that maximizes the width of syntactic categories. This deep parser uses the property grammar we have built according to the PEAS guide as an external resource. The behavior of all types of property is hardly coded in the parsers program. The interest of this deep parser is that it can be used to parse deeply, exhaustively and quite quickly, because of the inclusion of the constraint satisfaction process for property grammars directly in its kernel. Its drawback lies in the fact that it is necessary to modify the program if the behavior of a type of property needs to be changed.

LPL2 is a shallow deterministic parser. Its strategy relies on a left corner analysis, using in a first stage linearity and constituency constraints, before evaluating the entire grammar. This parser is very efficient (4 minutes for 1 million words). That kind of shallow parser is developed in a comparison perspective as well as for its ability to quickly give a surface analysis. Techniques used for this parser are very simple, which of course has an effect on the results obtained. The whole left context of a token in an analyzed sentence is taken into account, but only one token to the right. As LPL1, this shallow parser uses our EASy Property Grammar as an external resource. The behavior of properties is also hardly coded in the parsers program.

The third parser (LPL3) is a non-deterministic and deep parser as LPL1 developed in JAVA. For this parser, an XML version of the French grammar is used as an external resource. Two specific algorithms are used: one for parsing and one in order to produce a deterministic output compatible with the EASy evaluation procedure. The techniques used with LPL3 are very different than the ones used for LPL1. The constraint satisfaction process is helped while parsing by a filter (preventing too bad constructions from being kept) based on a measure of satisfaction density (SD.). The determinization is based too on categories that maximize contextually the measure of SD. Hereafter, this parser will also be called *Seed parser* because of its particular algorithm defined in (Vanrullen, 2005) which produces analysis by considering the grammar as a set of seeds able to react with the elements of the input. This parser uses the EASy Property Grammar as an external resource, as well as the behavior of properties which are defined in another XML file called the semantic specification of property grammars (see (Vanrullen, Guénot, Bellengier, 2003) for mode details). As for LPL1, the constraint satisfaction process for property grammars is directly programmed in its kernel, but the semantic specification of properties can be developed out of the program, by linguists and non programming persons. This supposes a metagrammar model, which allows many new types of properties to be introduced or modified without modifying the program. The drawback of this technique is that LPL3 is slower than LPL1 and needs further developments in order to preprocess and compile the semantic specification of constraints.

After the POS-tagging of the EASy corpora, each parser is tested, parameterized and finally used to produce its results. Note that the different parsing times of the three parsers allowed them more or less attempts: many for LPL2 (four minutes per attempt), three for LPL1 (one day per attempt), and only one for LPL3 (nearly four days per attempt).

Symbolic parsers have several assets contrary to stochastic ones, particularly if the formalism they use allows flexibility in their behavior; which is the case for Property Grammars with which the granularity of the analysis can be tuned. This tuning is done by choosing the type or the number of satisfied constraints needed to build a construction. With the same grammar and the same parsing strategy, it is possible to perform different analyzes giving more or less detailed results according to the needs.

5. Evaluation

One of the difficulties of the campaign lies in the diversity of the corpora and the their material. For an example, the oral, mail and question corpora contain, as expected, the highest proportion of non canonical inputs and represent 50% of the reference corpora.

The scores given for each corpus and for each parser are the classical precision, recall and f-score measures. Each value is calculated twice: once strictly (the reference and the parsers output should correspond boundary per boundary) and once vaguely (a tolerance of one boundary token per construction is allowed). The second kind of calculous gives what we will call here fuzzy precision, fuzzy recall and fuzzy f-score measures.

The types of syntactic categories evaluated are NV (verbal phrase), GN (noun phrase), GP (prepositional phrase), GA (adjectival phrase), GR (adverbial phrase) and PV (verbal phrase introduced by a preposition). Because the campaign is finishing, only some mean results are now known, and

not detailed ones, we only can discuss here some general facts about the EASy campaign. It is too early to describe particular data about the behavior of parsers category per category.

The batch of preliminary results which was calculated for our parsers is described hereunder in order to evaluate the relative quality of the parsers algorithms. Results are detailed per corpus type (oral, medical, general, literary, email) but not category per category. The scores are given parser per parser, with each time a strict and a fuzzy measure. A synthetic tabular of these scores is shown in the table 1 and figures 2, 3 and 4.

The table 1 shows the systematic superiority of LPL1 (deep parser) against both LPL2 and LPL3. Differences of fuzzy f-score between LPL1 and LPL3 are around 3% and 5% between LPL1 and LPL2. The second deep parser (LPL3) has better scores than the shallow parser (LPL2) excepted for medical and questions corpora.

Several conclusions come with this data:

Fuzzy scores are better than strict ones, which was foreseen. LPL1 gains about 4%, LPL2 about 8% and LPL3 also 4%. The shallow parser is thus less precise than the two others.

Parser	corpus type	precision	recall	f-score	fuzzy precision	fuzzy recall	fuzzy f-score
LPL1	general	85,16	85,17	85,94	87,35	87,46	88,06
	literature	84,41	86,75	85,53	86,91	89,62	88,14
	email	76,93	78,14	81,44	80,57	82,47	84,86
	medical	79,17	78,42	80,97	83,59	81,48	83,76
	oral	73,66	74,48	76,25	81,79	83,01	83,12
	questions	86,62	85,19	85,96	88,23	87,05	87,40

Parser	corpus type	precision	recall	f-score	fuzzy precision	fuzzy recall	fuzzy f-score
LPL2	general	73,14	78,29	77,55	76,44	82,04	80,77
	literature	68,64	75,53	72,25	74,46	82,52	77,98
	email	66,51	72,65	73,80	71,86	79,06	78,24
	medical	76,81	80,99	78,97	80,75	84,55	82,58
	oral	66,02	68,98	70,40	75,13	78,90	78,07
	questions	76,44	78,18	77,29	79,72	81,69	80,66

Parser	corpus type	precision	recall	f-score	fuzzy precision	fuzzy recall	fuzzy f-score
LPL3	general	80,80	80,71	81,54	83,50	83,47	84,14
	literature	81,95	83,94	82,89	84,72	87,07	85,78
	email	75,39	76,18	79,41	79,92	81,26	83,57
	medical	76,17	75,64	78,12	81,03	79,05	81,31
	oral	69,12	69,10	71,53	77,62	77,57	78,43
	questions	79,33	78,20	79,33	82,27	81,42	81,63

Table 1: Synthetic results for LPL1, LPL2 and LPL3

Deep parsing is globally better than shallow parsing, even for this evaluation framework which needs flat non hierarchical structures.

Literary and general texts are more grammatical than oral transcriptions and emails. It is thus easier to be guided by their structure to find a good construction. Such a tendency is well followed by parsers excepted for LPL2 whose score with the medical corpus is far better than with other corpora. LPL1 looses nearly 6% of quality between its best score (literary corpus) and its worst (oral). LPL2 looses 7% and LPL3 9%. These differences show how much deep parsing techniques are more sensible to the grammaticality of the corpora. However, due to the fact that constraints can be relaxed in the PG formalism, the differences stay acceptable.

The main differences between LPL1 and LPL3 vary between 2% to 4%. These parsers follow exactly the same variability according to the type of corpus. Their varia-

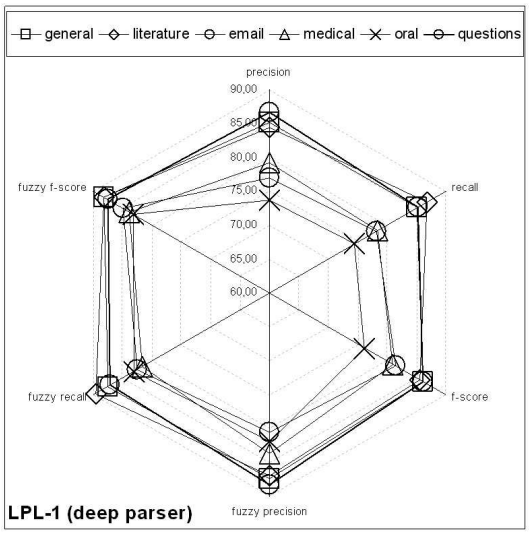

Figure 2: Strict and fuzzy scores for LPL-1

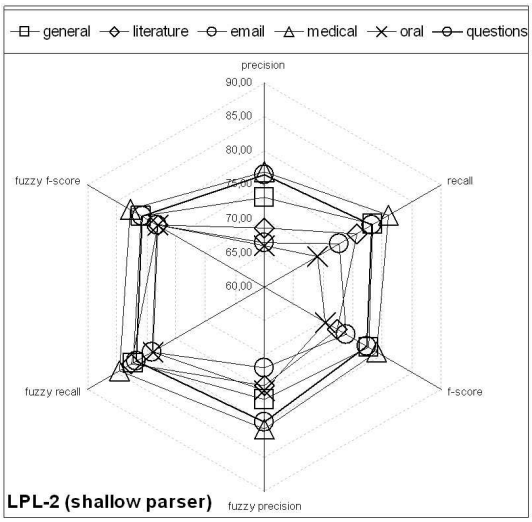

Figure 3: Strict and fuzzy scores for LPL-2

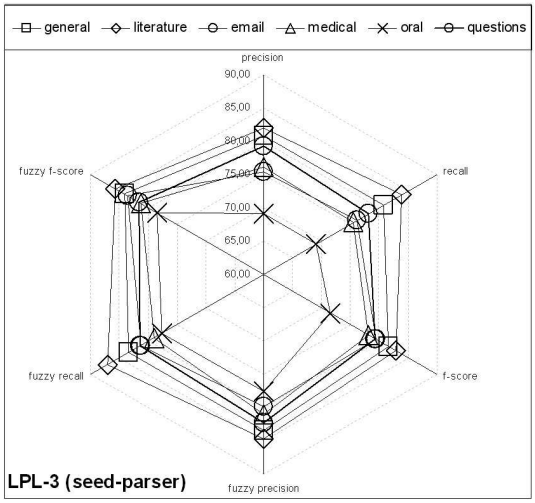

Figure 4: Strict and fuzzy scores for LPL-3

tions are practically parallel. The difference in sensitivity between these two analyzers can be explained because two lightly different grammars were used (in this case, the description of the differences will make it possible to refine grammars). Another explanation to this constant variation comes from the technique of determinization employed. This one does not belong to the PG formalism, but constitutes the only way to provide a single output starting from a non deterministic analysis. When the totality of the results will be available, it will be then possible to clearly highlight the various drawbacks of these techniques.

Figures 2, 3 and 4 allow the comparison of a given parser's scores corpus type per corpus type. They allow also a quick comparison of parsers one with another. It appears that the quality of the analyzes is better on normative corpora (literary, medical, questions) and worst on oral transcriptions. Emails and general corpora lead to mean results.

One note also a constancy for each parser between its scores of precision, recall and F-score, which encourages us to analyze more precisely their results concentrating only on F-score in order to apprehend the observable phenomena more easily. Table 2 and figure 5 give a clearer snapshot of the differences between parsers per corpus type.

corpus type	f-score LPL-1	fuzzy f-score LPL-1	f-score LPL-2	fuzzy f-score LPL-2	f-score LPL-3	fuzzy f-score LPL-3
general	85,94	88,06	77,55	80,77	81,54	84,14
literature	85,53	88,14	72,25	77,98	82,89	85,78
email	81,44	84,86	73,80	78,24	79,41	83,57
medical	80,97	83,76	78,97	82,58	78,12	81,31
oral	76,25	83,12	70,40	78,07	71,53	78,43
questions	85,96	87,40	77,29	80,66	79,33	81,63
Mean per type	**82,68**	**85,89**	**75,04**	**79,72**	**78,81**	**82,48**
Weighted mean	**80,29**	**84,84**	**73,51**	**79,27**	**76,21**	**81,03**

Table 2: F-scores per corpus type

The parsers show different results corresponding with what we could foresee: the shallow parser (LPL2) is globally less efficient than the two others. Mean f-scores for LPL1, LPL2 and LPL3 are respectively 84.8, 79.3 and 81 showing a clear correlation with the techniques and the strategies. More precisely, differences in parsing and determinization techniques can significantly explain the different scores. Even if grammar and lexicon were the same for the three parsers, their impact should not be forgotten: for example, deep parsing techniques can overcome tagging errors, which is not the case for the shallow parser. In our experiments, the pos-tagger performance was less than 90%: improving the pos-tagger will obviously improve the parsers. One interesting result is that there is a good stability for the three parsers of the results from one corpus type to another: only 5 points separate the literary and general corpora from the oral and mail ones. This is a clear indication of the robustness of the approach.

With these results, we would like to make some remarks about the EASy campaign and the three parsers. First, about the POS-tagging and the way to produce a non ambiguous parse: in some cases, the morphosyntactic features given by the POS-tagger are not precise enough to allow the selection of the good syntactic category: for an example it is difficult to choose between a GP and a GN introduced by a determiner amalgamated with a preposition like *des* (de + les).

Next, about the grammar. The PEAS guide gives different construction for a same morphosyntactic context depending on the membership of the words of the context to the French language. This depends too highly on the lexicon. Human annotators of the reference corpora and POS-taggers will not easily give the same interpretation of a foreign word.

Finally, about the PEAS guide imprecisions. In this guide, nothing is decided about repetitions and other phenomenons encountered with oral transcriptions. For repeated contiguous words, we chose to keep the first option consisting in the inclusion of repeated words into the same construction. Some human annotators may have chosen another possibility consisting in excluding repeated words from the construction. The example below shows the two possible hypothesis for a sentence extracted from an oral corpus:

- <**NV**> **il il se tachait** </**NV**> sa sa <**NV**> **il ne ne buvait** </**NV**> que des Blancs
- il <**NV**> **il se tachait** </**NV**> sa sa il ne <**NV**> **ne buvait** </**NV**> que des Blancs

All these remarks show that we must considerate the whole parsing process much that parsers themselves to understand and interpret the scores. A preliminary set of mean results for all 14 participants of the campaign indicates that the parsers LPL1,2 and 3 are respectively ranked at the third, sixth and fourth positions. Constraint based symbolic techniques prove they can have as good results as numerical ones. Moreover, symbolic shallow parsing can be very efficient both in terms of robustness, precision and runtime.

6. Conclusion

Several remarks come with this evaluation. First, the quality of the results indicates it becomes possible to consider automatic bracketing for large corpora, then the possibility of building large treebanks for French, including for spoken language corpora. This is clearly an important result and will make it possible to build a new kind of syntactic resource. The second important result is that symbolic techniques obtain very good results (including efficiency) in dealing with large corpora. The classical interests of symbolic techniques in terms of grammar development, adaptability and reusability come then on top of efficiency for such methods. We have shown in particular the possibility of reusing the same parsing mechanism (constraint satisfaction) and the same resources (grammar and lexicon) both for deep and shallow parsing, obtaining in both cases very good results. This also constitutes a technical validation of the theoretical framework, Property Grammars.

7. References

S. Abney. (1991). Parsing by chunks. In Berwick, R., Abney, S., Tenny, C. (Eds.). *Principle-based parsing.* Kluwer Academic Publishers, Dordrecht: 257-278.

S. Abney. (1997). Part-of-speech tagging and partial parsing. In Young, S., Bloothooft, G. *Corpus-Based Methods in Language and Speech Processing*, Kluwer Academic Publishers, Dordrecht: 118-136.

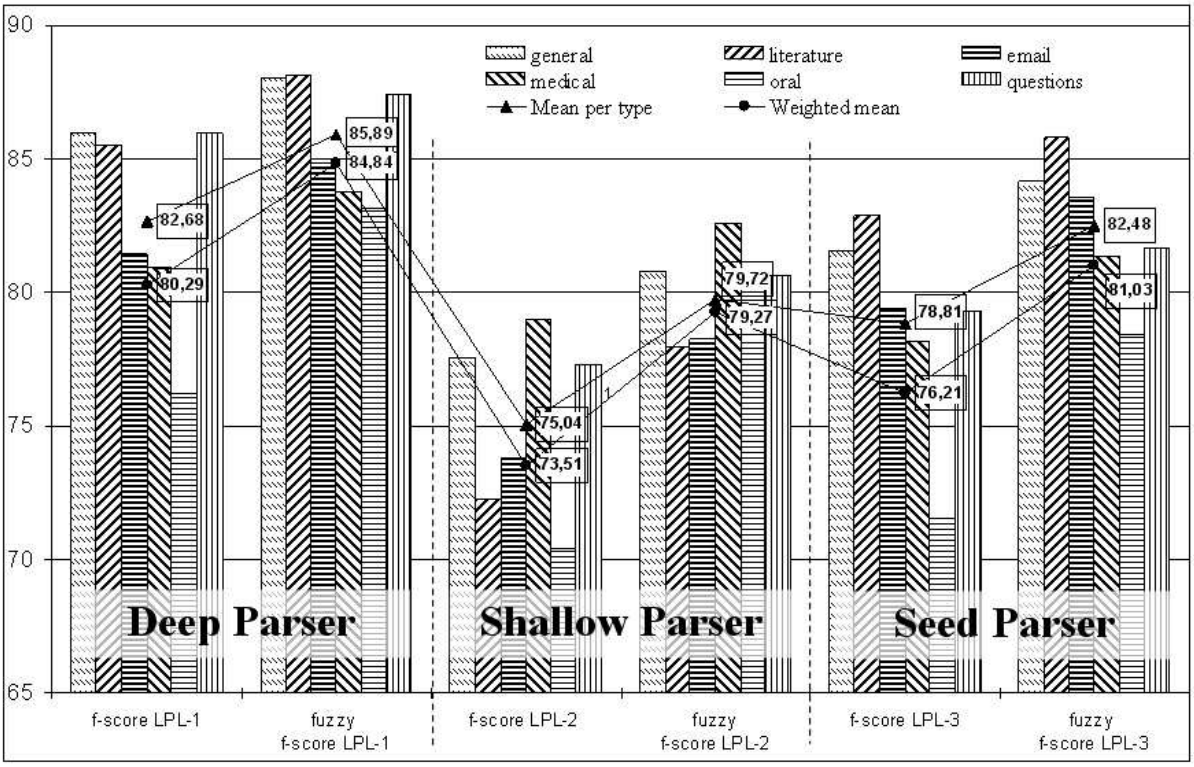

Figure 5: F-scores per corpus type

G. Bès. (1999). La phrase verbale noyau en franais, in *Recherches sur le franais parlé*, 15, Universit de Provence.

P. Blache. (2000). Constraints, Linguistic Theories and Natural Language Processing, in *Natural Language Processing*, D. Christodoulakis (ed), LNAI 1835, Springer-Verlag

P. Blache. (2001). *Les Grammaires de Propriétés : Des contraintes pour le traitement automatique des langues naturelles*, Hermès.

P. Blache, M.L. Guénot, T. Vanrullen. (2003). Corpus-based grammar development, in *proceedings of Corpus Linguistics-03*.

P. Blache. (2005). Property grammars : A fully constraint-based theory. In H Christiansen, P Skadhauge, J Villadsen, editors, *Constraint Satisfaction and Language Processing*. Springer-Verlag.

V. Gendner, G. Illouz, M. Jardino, L. Monceaux, P. Paroubek, I. Robba, A. Vilnat. (2002). A Protocol for Evaluating Analyzers of Syntax (PEAS) , *Proceedings of LREC-2002*, Las Palmas de Gran Canaria, Spain, 27 May- 2 June 2002.

V. Gendner, G. Illouz, M. Jardino, L. Monceaux, P. Paroubek, I. Robba, A. Vilnat. (2003). PEAS, the fi rst instantiation of a comparative framework for evaluating parsers of French, Research notes of the *10th Conference of the European Chapter of the Association for Computional Linguistics, EACL03* ,Budapest, April 2003.

V. Gendner, A. Vilnat. (2004). *Les annotations syntaxiques de référence PEAS*, version 1.6. Révisions par : Laura Monceaux, Patrick Paroubek, Isabelle Robba.

J. Hirschberg, O. Rambow. (2001). Learning Prosodic Features using a Tree Representation. AT&T Labs Research. *Eurospeech 2001* - Scandinavia.

M. Liberman, K. Church. (1992). Text analysis and word pronunciation in text-to-speech synthesis. In Furui, S., Sondhi, M.M. (Eds), *Advances in Speech Signal Processing*, New York: Dekker, 791-831.

T. Vanrullen, P. Blache. (2002). An evaluation of different symbolic shallow parsing techniques, in *proceedings of LREC-02*.

T. Vanrullen, Marie-Laure Guénot, Emmanuel Bellengier. (2003). Formal representation of property grammars. In *Proceedings of ESSLLI Student Session*.

T. Vanrullen. (2005). *Vers une analyse syntaxique à granularité variable*. Thèse, Université de Provence.

T. Vanrullen, P. Blache, C. Portes, S. Rauzy, J.F. Maeyhieux, J.M. Balfourier, M.L. Guénot, E. Bellengier. (2005). Une plateforme pour lacquisition, la maintenance et la validation de ressources lexicales. in *proceedings of TALN 2005*.

A. Vilnat, P. Paroubek, L. Monceaux, I. Robba, V. Gendner, G. Illouz, M. Jardino. (2003). EASY or How difi cult can it be to defi ne a reference corpus for French?, *Proceedings of the second Workshop on TreeBanks ans Linguistic Theories, TLT 2003*,Växjö, Sweden, November 2003.

A. Vilnat, P. Paroubek, L. Monceaux, I. Robba, V. Gendner, G. Illouz, M. Jardino. (2004). The ongoing Evaluation Campaign of Syntactic Parsing of French: EASY, *LREC 2004*, Lisboa, Portugal, May 2004.