Using hybrid probabilistic-linguistic knowledge to improve pos-tagging performance[1]

Montserrat Civit Torruella (*)
Mª Antònia Martí Antonín (**)
Lluís Padró Cirera (*)

(*) TALP Research Centre
Universitat Politècnica de Catalunya
{civit,padro}@talp.upc.es
(**) CLiC Centre de Llenguatge i Computació
Universitat de Barcelona
amarti@fil.ub.es

Abstract

We present the tuning of a statistical PoS tagger via the inclusion of hand-written linguistically motivated context constraints.

The used tagger is able to use information of any degree: n-grams, automatically learned context constraints, linguistically motivated manually written constraints, etc. The sources and kinds of constraints are unrestricted, and the language model can be easily extended, improving the results. The tagger has been trained, tuned and tested using a high quality, hand-checked, 100,000-token Spanish corpus.

Obtained results show that, although the inclusion of hand-written context rules in the tagger model only raise the tagger precision from 97.3% to 97.4% at category-subcategory level, the precision at a more detailed level (considering morphological features such as number, gender, person, etc.) is raised from 94.5% to 96.7%.

1. Introduction

In NLP, it is necessary to model the language in a representation suitable for the task to be performed. The language models more commonly used are based on two main approaches. First, the linguistic approach, in which the model is written by a linguist, generally in the form of rules or constraints (Voutilainen & Järvinen 95). Second, the automatic approach, in which the model is automatically obtained from corpora (either raw or annotated)[2], and consists of n-grams (Garside et al. 87, Cutting et al. 92), rules (Hindle 89) or neural nets (Schmid 94).

In the automatic approach we can distinguish two main trends. The low-level data trend collects statistics from the training corpora in the form of n-grams, probabilities, weights, etc. The high-level data trend acquires more sophisticated information, such as context rules, constraints, or decision trees (Daelemans et al. 96, Màrquez & Rodríguez 95, Samuelsson et al. 96). Still another possibility is hybrid models, which try to join the advantages of both approaches (Voutilainen & Padró 97) as the one presented here.

We present in this paper a hybrid approach that puts together both trends, automatic and linguistic approach. We describe a POS tagger based on the work described in (Padró 96, Màrquez & Padró 97), which is able to use bi/trigram information, automatically learned context constraints and linguistically motivated manually written constraints. The sources and kinds of constraints are unrestricted, and the

[1]This research has been partially funded by the Spanish Research Department (CICyT TIC2000-1735-C02-02, TIC2000-0335-C03-02; X-Tract-II BFF2002-04226-C03-03), and by the UE Comission (IST-2001-34460).
[2] When the model is obtained from annotated corpora we talk about supervised learning, when it is obtained from raw corpora training is considered unsupervised.

language model can be easily extended. In this paper we focus on the collaboration of simple statistical constraints (bigrams) and a small set of accurate hand-written rules.

The paper is organized as follows: in section 2 we present the tagging algorithm and the kind of language model used. In section 3, we describe the linguistic constraints. Finally, descriptions of the corpus used, the experiments performed and the results obtained can be found in section 4.

2. Algorithm and language model

2.1 Language model

The constraint language used is able to express the same kind of patterns than the Constraint Grammar formalism (Karlsson et al. 95), although the formalism has been extended to allow each constraint to have a compatibility value that indicates its strength.

We use a hybrid language model consisting of an automatically acquired part and a linguist-written part. The automatically acquired part consists of bigrams collected from the annotated training corpus. The linguistic part is comparatively small, and aims to provide high precision results for the cases that are not accurately captured by the statistical information.

A sample rule of the linguistic part is:

```
15.1 (<P*>)
   (0 (<D*>))
   (1 (<VMI*>) OR (<VMS*>));
```

This rule states that the tag P (*pronoun*) is more compatible than D (*determiner*) with a following word having as a tag VMI (main verb indicative) or VMS (subjunctive). The number in the first line (15.1) indicates the compatibility degree, while numbers in other lines indicate the position of the referred elements: 0 stands for the same position (it indicates the ambiguity class); 1 indicates the left position; -1 the right position, and so on. The *-symbol indicates that the rule applies for all the tags starting by P, D, VMI or VMS, whatever other values they have (i.e. P* applies for PP3FSA00 -feminine singular third person accusative personal pronoun- as well as for PD0MS000 -masculine singular demonstrative pronoun-; D* for DA0FP0 -feminine plural definite article- as well as for DI0MS0 -masculine singular indefinite determiner-, etc.). The formalism also allows the expression of disjunctions and negative conditions. In the preceding rule, there is a disjunction when indicating the left context [(VMI*) OR (VMS*)]. In the following rule, there is a negative condition (NOT 1 "ser"):

```
1.5 (<PP3MSA00>)
   (0 (<PP3CNA00>))
   (NOT 1 "ser");
```

which gives more weight to the PP3MSA00 tag if the following word does not have the lemma "ser" ('to be'). As we can see, constraints not only deal with PoS-tags, but also with lemmas, even with words, as we will show later is this paper.

2.2 Tagging algorithm

Usual tagging algorithms are either n-gram oriented -such as Viterbi algorithm (Viterbi 67)- or ad-hoc for every case when they must deal with more complex information. We use relaxation labelling as a tagging algorithm. Relaxation labelling is a generic name for a family of iterative algorithms which perform function optimisation, based on local information -see (Torras 89) for a summary-. Its most remarkable feature is that it can deal with any kind of constraints, thus the model can be improved by adding any constraints available and it makes the tagging algorithm independent of the complexity of the model.

The algorithm has been applied to part-of-speech tagging (Padró 96), shallow parsing (Voutilainen 97 & Padró 97), semantic parsing (Atserias et al. 01), and other NLP tasks (Daudé et al. 00).

The algorithm is described as follows:

Let $V=\{v_1,v_2,...,v_n\}$ be a set of variables (words).
Let $t_i=\{t_1^i,t_2^i,...,t_{mi}^i\}$ be the set of possible labels (POS tags) for variable v_i.

Let CS be a set of constraints between the labels of the variables. Each constraint C in CS states a "compatibility value" C_r for a combination of pairs variable-label. Any number of variables may be involved in a constraint.

The aim of the algorithm is to find a weighted labelling[3] such that "global consistency" is maximized. Maximizing "global consistency" is defined as maximizing for all v_i, $\sum_j p^i_j \times S_{ij}$, where p^i_j is the weight for label j in variable v_i and S_{ij} the support received by the same combination. The support for the pair variable-label expresses *how compatible* that pair is with the labels of neighbouring variables, according to the constraint set. It is a vector optimisation and does not maximize *only* the sum of the supports of all variables. It finds a weighted labelling such that any other choice would not increase the support for *any* variable.

The support is defined as the sum of the influence of every constraint on a label, $S_{ij} = \sum_{r \in R_{ij}} Inf(r)$ where:

- R_{ij} is the set of constraints on label j for variable i, i.e. the constraints formed by any combination of variable-label pairs that includes the pair (v_i, t^i_j).
- $Inf(r) = C_r \times p^{r1}_{k1}(m) \times ? \times p^{rd}_{kd}(m)$, is the product of the current weights[4] for the labels appearing in the constraint except (v_i, t^i_j) (representing *how applicable* the constraint is in the current context) multiplied by C_r which is the constraint compatibility value (stating *how compatible* the pair is with the context).

Briefly, what the algorithm does is:

1. Start with a random weight assignment (We use lexical probabilities as a starting point).
2. Compute the support value for each label of each variable.
3. Increase the weights of the labels more compatible with the context (support greater than 0) and decrease those of the less compatible labels (support less than 0, negative values for support indicate *incompatibility*), using the updating function:
$$p^i_j(m+1) = p^i_j(m) \times (1+ S_{ij}) \;/\; \sum_{k=1,ki} p^i_k(m) \times (1+ S_{ik}) \quad \text{where} \quad -1 \le S_{ij} \le +1$$
4. If a stopping/convergence criterion is satisfied, stop, otherwise go to step 2. We use the criterion of stopping when there are no more changes, although more sophisticated heuristic procedures are also used to stop relaxation processes (Eklundh & Rosenfeld 78, Richards et al. 81).

The cost of the algorithm is proportional to the product of the number of words by the number of constraints.

3. Linguistic rules

Hand-written rules may concern the so-called *short*-tag (i.e. the first two elements of the tag indicating the morphological category, (for instance, pronoun) and its subclass (personal, demonstrative, etc)) or the *long*-tag, including all the morphological category features (gender and number for nouns; time, person, and number for verbs, etc.)[5].

In order to establish the hand-written constraints, the main errors commited by the tagger when automatically annotating the CLiC-TALP corpus (see next section) were studied. The most frequent errors (ten or more occurrences) appear in table 1, in which, the number of occurrences (*#occ*) is shown for each kind of error, as well as the involved word, the tag proposed by the tagger, and the correct tag according to the reference corpus.

[3]A weighted labelling is a weight assignment for each label of each variable such that the weights for the labels of the same variable add up to one.
[4]p^r_k (m) is the weight assigned to label k for variable r at time m.
[5] See section 4.1 and (Civit 00) for details about the tagset.

#occ	Word	Proposed	Expected	#occ	Word	Proposed	Expected
411	se	P0000000	P0300000	15	orden	NCFS000	NCMS000
315	que	PR0CN000	CS	14	habría	VAIC1S0	VAIC3S0
152	lo	PP3CNA00	PP3MSA00	13	haya	VASP1S0	VASP3S0
127	se	P0000000	PP3CN000	13	mucho	RG	DI0MS0
115	era	VSII1S0	VSII3S0	13	sí	PP3CNO00	RG
80	que	CS	PR0CN000	13	me	PP1CS000	P010S000
78	había	VSII1S0	VSII3S0	13	parecía	VSII1S0	VSII3S0
39	tenía	VSII1S0	VSII3S0	12	llamaba	VSII1S0	VSII3S0
34	estaba	VSII1S0	VSII3S0	12	una	DI0FS0	DN0FS0
30	le	PP3CSD00	PP3CSA00	12	poco	RG	PI0MS000
29	Se	P0000000	P0300000	11	daba	VSII1S0	VSII3S0
26	la	DA0FS0	PP3FSA00	11	final	NCFS000	NCMS000
24	sea	VSSP1S0	VSSP3S0	11	defensa	NCCS000	NCFS000
24	sería	VSIC1S0	VSIC3S0	10	bueno	AQ0MS0	I
22	sí	RG	PP3CNO00	10	los	DA0MP0	PP3MPA00
21	podía	VSII1S0	VSII3S0	10	quería	VSII1S0	VSII3S0
20	podría	VMIC1S0	VMIC3S0	10	fuera	VSSI1S0	VSSI3S0
19	hacía	VSII1S0	VSII3S0	10	nada	PI0CS000	RG
19	hubiera	VASI1S0	VASI3S0	10	sabía	VSII1S0	VSII3S0
16	decía	VSII1S0	VSII3S0	10	un	DI0MS0	DN0MS0
15	Era	VSII1S0	VSII3S0	10	una	DI0FS0	PI0FS000
15	mismo	RG	DI0MS0				

Table 1: most frequent errors after the automatic tagging

The main source of errors concerns words *que* and *se*. The first one may receive only two tags: relative pronoun or subordinating conjunction, but contexts in which it may appear are the same, as shown in the following examples, in which *que* is a conjunction in (a) and a relative pronoun in (b).

<*adjective* + *'que'*>
(a) Es *probable que* sea más *fácil que* la segunda célula se dispare[6].
(b) El escándalo *público que* provocó aquella decisión[7].
<*preposition* + *'que'*>
(a) Se ha deslizado en la mente de los españoles la convicción *de que* no somos refinados[8].
(b) Esa dignidad en el comportamiento público *de que* hacéis gala[9].
<*noun* + *'que'*>
(a) Druso ordenó a la *tropa que* plantara en la cima una bandera[10].
(b) Por muy intensa que sea la *escena que* se represente[11].
<*verb* + *'que'*>
(a) Y yo queriendo hacer *ver que* no podían notarme nada[12].
(b) Las veredas sin *urbanizar que* habían quedado abiertas entre las chozas[13].
<*adverb* + *'que'*>
(a) Es todavía hoy un aviso *más que* una constatación[14].
(b) Brindar soluciones globales impuestas desde *arriba que* la hagan superflua o dañina[15].

[6] It will probably be easier that the second cell fires.
[7] The public scandal that caused that decision.
[8] It has slipped into the minds of Spanish people the belief that we are not refined.
[9] That dignity you display in your public behaviour.
[10] Drusus ordered the troops to plant a flag at the top.
[11] However intense the scene that will be performed is.
[12] And there I was, trying to make believe that they could not notice anything about me.
[13] The unurbanised paths that were left opened between the huts.
[14] It is still nowadays more of a warning than a verification.
[15] She affords global solutions imposed from above that will make it superfluous or harmful.

As for the word *se*, its disambiguation depends on the use of the verb. The three tags it may receive correspond to a pronominal form, a sentence mark or a verbal mark. The first one, pronoun (PP3CN000), is used when *se* has a syntactic function as direct or indirect object; the second one (P0000000) is used when it marks that the sentence is an impersonal or a passive one; finally, the last tag (P0300000) is used to mark middle voice, pronominal verb, emphatic value, etc.

In both cases (*que* and *se*), the disambiguation requires more than purely formal knowledge. For *que*, the needed information is related to verb argument structure. For *se*, information about verbs use and meaning is needed in order to disambiguate. Since we do not have such information, constraints were only introduced when the formal, local context provided enough information.

Hand-written constraints aim to cover as many errors as possible. Therefore we focused on the previously shown cases (cf. table 1).
Rules concerning the short tag are the following:

```
R1: 15.1 (<P*>)
         (0 (<D*>)
         (1 (<VMI*>) OR (<VMS*>) OR (<VAI*>) OR (<VAS*>) OR (<VSI*>)
         OR (<VSS>));
```

R1 states that pronominal forms (P*) are more compatible before verbal forms than determiners (D*).

```
R2:  10.1 (<PP3*>)                 10.2 (<RG>)
          (0 (<RG>))                    (0 (<PP3*>))
          (-1 (<SPS00>);           (NOT -11 (<SPS00>));
```

R2 gives more weight to the pronominal tag when it comes after a preposition (SPS00), and to the adverbial tag (RG) if the preceding word is not a preposition.

```
R3: 5.0 (<CS>)
        (0 (<>PR*)
        (-1 ("de"))
        (-2 (<NC*>);
```

R3 states that after a sequence of the preposition *de* and a noun (NC), tag CS (conjunction) is preferable instead of the relative pronoun tag (PR*). This constraint concerns the word *que*.

```
R4: 5.0 (<NC*>)
        (0 (<RG>))
        (-1 (<D*>));
```

R4 applies for some words that may receive adverbial (RG) and nominal (NC*) tags, and selects NC if the previous word is a determiner.

Rules concerning long tags are shown in what follows.

```
R5: 0.5 (<VMSP3S0>)
        (0 (<VMSP1S0>));
```

R5 is only a sample of a set of twelve rules concerning the verbal person. They give a higher weight to the third person, because it is by far the most frequent in the corpus (8685 verbal forms of the third person versus 896 of the first one). Verbal tenses concerned by this rule are all imperfect forms, the conditional and the present of subjunctive, and they apply for main, auxiliary and semi auxiliary verbs.

```
R6: 1.5 (<PP3CNA00>)              1.0 (<PP3MSA00>)
        (0 (<PP3MSA00>))              (0 (<PP3CNA00>))
        (1 ("ser"));                  (NOT 1 ("ser"));
```

R6 applies for the word *lo*, which may receive two tags when being a pronoun: masculine singular or common invariable. Whether one or another tag is correct depends on the following verbal form. If the verb is a copulative one, the correct tag is the common-invariable; otherwise, the masculine-singular. Since there are three copulative verbs in Spanish and they may appear as simple as well as complex forms, there are twelve more rules concerning this ambiguity class.

```
R7: 5.0 (<P0*>)
        (0 (<"-se-">))
        (1 ("atreverse") OR ("abrumarse") OR ("adjudicarse") ... );
```

R7 concerns the word *se* when appearing with the so-called 'pronominal verbs'. Pronominal verbs are those having the suffix *-se* in the infinitive. When they have a finite form, *se* always precedes them but it does not have a special value. As *se* has person variation (*me, te, nos, os*), more rules were introduced in order to account for this phenomenon. This is the only case in which it is possible to know a priori the value of this word.

R8: 1.1 (<NCM*>)
(0 (<NCF*>))
(-1 (<DN0M*>)
OR (<DD0M*>) ...);

1.1 (<NCF*>)
(0 (<NCM*>))
(-1(<DN0F*>)
OR (<DD0M*>)...);

R8 selects the tag for the nominal gender (masculine or feminine) according to this value in the tag of the previous determiner. The position -1 contains the list of all kinds of determiners.

The last hand-written restriction concerns verbal lemmas. Spanish verbs are sorted out into three groups according to the end of the infinitive. The first group contains all verbs whose infinitive finishes by *-ar*. It is by far the biggest group and almost all forms are regular (it contains 11386 forms in our morphological lexicon). It is also the group in which new verbal forms are included. Second group contains verbs finishing by *–er* (609 forms); and the third group those ending by *–ir* (660 forms). Second and third group contain irregular verbs. The number of irregular verbs in Spanish is low and if they exist that means that they are very frequent. So, when a conjugated verbal form may belong to the regular group or to the irregular ones, it usually belongs to the latter. Thus, several restrictions concerning the choice of the lemma were introduced, in which the irregular form is preferred:

5.0 ("salir")
("salgar");

5.0 ("ser")
("erar");

There are twelve restrictions such as these ones. The improvement they give to the tagger has not been calculated because there were few erroneous cases in the corpus. However, they are really qualitatively useful in the treatment of huge amounts of text.

4. Experiments

4.1 Description of the corpus

We use the CLiC-TALP corpus, a 100,000 token Spanish corpus, developed by the CLiC-TALP[16] group. A 30,000 words subset is reserved for testing. The remaining 70,000 tokens are used as training and tuning material.

The CLiC-TALP corpus is a balanced subset of a larger one: Lexesp (Sebastián et al. 00), including originally written Spanish from both Spain and South-America. Texts are extracts from novels, scientific and weekly magazines, newspapers and sports papers ranging from 1978 to 1995.

The tagset used follows EAGLES recommendations (Monachini & Calzolari 96). The amount of tags is 285. Main categories are 13: noun, verb, adjective, adverb, pronoun, determiner, conjunction, preposition, interjection, dates, numbers, abbreviations and punctuation marks. Morphological features include gender, number, person, tense and mood. The tagging of the corpus was manually validated after an automatic tagging process, so CLiC-TALP corpus constitutes a reference corpus for Spanish.

4.2 Experiments and results

We start from a baseline tagger that uses only bigram information acquired from the training corpus. The performance of the tagging presented here is tuned over the same training corpus via the inclusion of hand-written linguistic constraints. Linguistic constraints are included in the model one by one in order to test its improvement.

In this section we present the results obtained by each rule. We report its precision both at a coarse level (category and subcategory: short tag) and at a fine-grained level (category, subcategory and morphological features: long tag).

[16] http://www.talp.upc.es, http://clic.fil.ub.es

	precision (coarse)	precision (fine)	precision (fine+lemma)
Bigrams	97.29	94.48	94.36
Bigrams +R1	**97.34**	94.53	94.41
Bigrams +R2	**97.34**	94.53	94.41
Bigrams +R3	97.30	94.50	94.37
Bigrams +R4	97.29	94.49	94.36
Bigrams +R5	97.29	**95.89**	**95.76**
Bigrams +R6	(97.34)	**94.74**	**94.62**
Bigrams +R7	97.29	94.48	94.36
Bigrams +R8	97.29	94.56	94.44
TOTAL	**97.40**	**96.66**	**96.18**

These results show that the improvement provided by the manual constraints at a coarse level is not significant, but that there is a large precision increase when the evaluation is performed at a fine-grained level. This points that the bigram information adequately captures the category-subcategory information, and that the errors made at this level are not easy to solve using only morphosyntactic information.

Some detailed comments need to be done. R6 depends on R1: the distinction between determiner and pronoun is previous to the distinction between the masculine or neuter form of pronouns; that is why performance of R6 was calculated together with performance of R1. The most significant hand-written rule is R5, because verbal forms are highly frequent in the corpus, especially the *había* form, which appears in complex verb forms.

After this phase, the remaining errors (ten or more) are shown in table 2:

#occ	Word	Proposed	Expected
411	se	P0000000	P0300000
313	que	PR0CN000	CS
127	se	P0000000	PP3CN000
82	que	CS	PR0CN000
30	le	PP3CSD00	PP3CSA00
29	Se	P0000000	P0300000
19	lo	PP3MSA00	PP3CNA00
15	mismo	RG	DI0MS0
13	me	PP1CS000	P010S000

#occ	Word	Proposed	Expected
13	mucho	RG	DI0MS0
12	poco	RG	PI0MS000
12	una	DI0FS0	DN0FS0
11	defensa	NCCS000	NCFS000
10	buen	AQ0MS0	I
10	nada	PI0CS000	RG
10	un	DI0MS0	DN0MS0
10	una	DI0FS0	PI0FS000

Table 2: remaining errors after hand-written rules

5. Conclusions and further work

We have presented a PoS tagging model for unrestricted Spanish text, consisting of a statistical bigram-based model, enhanced with a reduced set of hand-written context constraints. The manual constraints are incrementally build and focus on the main errors made by the statistical model.

Performed experiments yield that hand-written rules significantly improve the results of the purely statistical tagger, especially at the fine level. It is noteworthy that with a few set of new restrictions, the increase in precision is considerable.

Further work should include the coverage of remaining errors, though most of the unsolved cases require semantic information, which is not available at this stage of the process. To make it possible, the constraint formalism should be extended to deal with semantic context features.

References

Atserias J, Padró L, Rigau G 2001 Integrating Multiple Knowledge Sources for Robust Semantic Parsing. Proceedings of the International Conference on *Recent Advances on Natural Language Processing* RANLP'01. Tzigov Chark, Bulgaria. 2001.

Civit M 2000 Guía para la anotación morfológica de corpus *X-Tract WP-00/06*. Available at http://www.lsi.upc.es/civit/publicacions.html

Cutting D, Kupiec J, Pederson J, Sibun P 1992 A Practical Part-of-Speech Tagger. In *Proceedings of the Third Conference on Applied Natural Language Processing.*, ACL.

Daudé J, Padró L, Rigau G 2000 Mapping WordNets Using Structural Information *38th Annual Meeting of the Association for Computational Linguistics (ACL'2000)*, Hong Kong, October 2000.

Eklundh J, Rosenfeld A 1978 Convergence Properties of Relaxation Labelling. Technical Report no. 701. Computer Science Center. University of Maryland.

Daelemans W, Zavrel J, Berck P, Gillis S 1996 MTB: A Memory-Based Part-of-Speech Tagger Generator. In *Proceedings of 4th Workshop on Very Large Corpora.* Copenhagen, Denmark.

Garside R, Leech G, Sampson G (eds) 1987 *The Computational Analysis of English.* London and New York: Longman.

Hindle D 1989 Acquiring disambiguation rules from text. In *Proceedings ACL'89*.

Karlsson F, Voutilainen A, Heikkila J, Anttila A (eds) 1995 *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text.* Mouton de Gruyter, Berlin and New York.

Màrquez L, Rodríguez H 1995 Towards Learning a Constraint Grammar from Annotated Corpora Using Decision Trees. ESPRIT BRA-7315 Acquilex II, Working Paper.

Monachini M, Calzolari N 1996 *Synopsis and Comparison of Morphological Phenomena Encoded in Lexicons and Corpora. A common Proposal and Applications to European Languages*, Eagles.

Padró L 1996 POS Tagging Using Relaxation Labelling. In *Proceedings of 16th International Conference on Computational Linguistics.* Copenhagen, Denmark.

Richards J, Landgrebe D, Swain P 1981 On the accuracy of pixel relaxation labelling. *IEEE Transactions on System, Man and Cybernetics.* Vol. SMC-11

Samuelsson C, Tapanainen P, Voutilainen A 1996 Inducing Constraint Grammars. In *Proceedings of the 3rd International Colloquium on Grammatical Inference.*

Schmid H 1994 Part-of-speech tagging with neural networks. In *Proceedings of 15th International Conference on Computational Linguistics.* Kyoto, Japan.

Sebastián N, Martí MA, Carreiras MF, Cuetos F 2000 *LEXESP: Léxico Informatizado del Español* Edicions de la Universitat de Barcelona.

Torras C 1989 Relaxation and Neural Learning: Points of Convergence and Divergence. *Journal of Parallel and Distributed Computing.* 6:217-244

Viterbi A J 1967 Error bounds for convolutional codes and an asymptotically optimal decoding algorithm. In *IEEE Transactions on Information Theory.* pg 260--269, April.

Voutilainen A, Järvinen T 1995 Specifying a shallow grammatical representation for parsing purposes. In *Proceedings of the 7th meeting of the European Association for Computational Linguistics.* 210-214.

Voutilainen A, Padró L 1997 Developing a Hybrid NP Parser. In *Proceedings of ANLP'97*.