Mapping Multilingual Hierarchies Using Relaxation Labeling

J. Daudé, L. Padró & G. Rigau
TALP Research Center
Departament de Llenguatges i Sistemes Informàtics
Universitat Politècnica de Catalunya. Barcelona
{daude,padro,g.rigau}@lsi.upc.es

Abstract

This paper explores the automatic construction of a multilingual Lexical Knowledge Base from pre-existing lexical resources. We present a new and robust approach for linking already existing lexical/semantic hierarchies. We used a constraint satisfaction algorithm (relaxation labeling) to select –among all the candidate translations proposed by a bilingual dictionary– the right English WordNet synset for each sense in a taxonomy automatically derived from a Spanish monolingual dictionary. Although on average, there are 15 possible WordNet connections for each sense in the taxonomy, the method achieves an accuracy over 80%. Finally, we also propose several ways in which this technique could be applied to enrich and improve existing lexical databases.

1 Introduction

There is an increasing need of having available general, accurate and broad coverage multilingual lexical/semantic resources for developing NL applications. Thus, a very active field inside NL during the last years has been the fast development of generic language resources.

Several attempts have been performed to produce multilingual ontologies. In (Ageno et al., 1994), a Spanish/English bilingual dictionary is used to (semi)automatically link Spanish and English taxonomies extracted from DGILE (Alvar, 1987) and LDOCE (Procter, 1987). Similarly, a simple automatic approach for linking Spanish taxonomies extracted from DGILE to WordNet (Miller et al., 1991) synsets is proposed in (Rigau et al., 1995). The work reported in (Knight and Luk, 1994) focuses on the construction of Sensus, a large knowledge base for supporting the Pangloss machine translation system. In (Okumura and Hovy, 1994) (semi)automatic methods for associating a Japanese lexicon to an English ontology using a bilingual dictionary are described. Several experiments aligning EDR and WordNet ontologies are described in (Utiyama and Hasida, 1997). Several lexical resources and techniques are combined in (Atserias et al., 1997) to map Spanish words from a bilingual dictionary to WordNet, and in (Farreres et al., 1998) the use of the taxonomic structure derived from a monolingual MRD is proposed as an aid to this mapping process.

This paper presents a novel approach for merging already existing hierarchies. The method has been applied to attach substantial fragments of the Spanish taxonomy derived from DGILE (Rigau et al., 1998) to the English WordNet using a bilingual dictionary for connecting both hierarchies.

This paper is organized as follows: In section 2 we describe the used technique (the relaxation labeling algorithm) and its application to hierarchy mapping. In section 3 we describe the constraints used in the relaxation process, and finally, after presenting some experiments and preliminary results, we offer some conclusions and outline further lines of research.

2 Application of Relaxation Labeling to NLP

Relaxation labeling (RL) is a generic name for a family of iterative algorithms which perform function optimization, based on local information. See (Torras, 1989) 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 the algorithm is independent of the complexity of the model. That is, we can use more sophisticated constraints without changing the algorithm.

The algorithm has been applied to POS tagging (Màrquez and Padró, 1997), shallow parsing (Voutilainen and Padró, 1997) and to word sense disambiguation (Padró, 1998).

Although other function optimization algorithms could have been used (e.g. genetic algorithms, simmulated annealing, etc.), we found RL to be suitable to our purposes, given its ability to use models based on context constraints, and the existence of previous work on applying it to NLP tasks.

Detailed explanation of the algorithm can be found in (Torras, 1989), while its application to NLP tasks, advantages and drawbacks are addressed in (Padró, 1998).

2.1 Algorithm Description

The Relaxation Labeling algorithm deals with a set of variables (which may represent words, synsets, etc.), each of which may take one among several different labels (POS tags, senses, MRD entries, etc.). There is also a set of constraints which state compatibility or incompatibility of a combination of pairs variable–label.

The aim of the algorithm is to find a weight assignment for each possible label for each variable, such that (a) the weights for the labels of the same variable add up to one, and (b) the weight assignation satisfies –to the maximum possible extent– the set of constraints.

Summarizing, the algorithm performs constraint satisfaction to solve a consistent labeling problem. The followed steps are:

1. Start with a random weight assignment.
2. Compute the *support* value for each label of each variable. Support is computed according to the constraint set and to the current weights for labels belonging to context variables.
3. Increase the weights of the labels more compatible with the context (larger support) and decrease those of the less compatible labels (smaller support). Weights are changed proportionally to the support received from the context.
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 may also be used to stop relaxation processes (Eklundh and Rosenfeld, 1978; Richards et al., 1981).

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

2.2 Application to taxonomy mapping

As described in previous sections, the problem we are dealing with is to map two taxonomies. That is:

- The starting point is a sense disambiguated Spanish taxonomy –automatically extracted from a monolingual dictionary (Rigau et al., 1998)–.
- We have a conceptual taxonomy (e.g. WordNet (Miller et al., 1991)), in which the nodes represent concepts, organized as *synsets*.
- We want to relate both taxonomies in order to have an assignation of each sense of the Spanish taxonomy to a WN synset.

The modeling of the problem is the following:

- Each sense in the Spanish taxonomy is a variable for the relaxation algorithm.
- The possible labels for that variable, are all the WN synsets which contain a word that is a possible translation of the Spanish sense. Thus, we will need a bilingual dictionary to know all the possible translations for a given Spanish word. This has the effect of losing the sense information we had in the Spanish taxonomy.
- The algorithm will need constraints stating whether a synset is a suitable assignment for a sense. These constraints will rely on the taxonomy structure. Details are given in section 3.

3 The Constraints

Constraints are used by relaxation labeling algorithm to increase or decrease the weight for a variable label. In our case, constraints increase the weights for the connections between a sense in the Spanish taxonomy and a WordNet synset. Increasing the weight for a connection implies decreasing the weights for all the other possible connections for the same node. To increase the weight for a connection, constraints look for already connected nodes that have the same relationships in both taxonomies.

Although there is a wide range of relationships between WordNet synsets which can be used to build constraints, we have focused on the hyper/hyponym relationships. That is, we increase the weight for a connection when the involved nodes have hypernyms/hyponyms also connected. We consider hyper/hyponym relationships either directly or indirectly (i.e. ancestors or descendants), depending on the kind of constraint used.

Figure 1 shows an example of possible connections between two taxonomies. Connection C_4 will have its weight increased due to C_5, C_6 and C_1, while connections C_2 and C_3 will have their weights decreased.

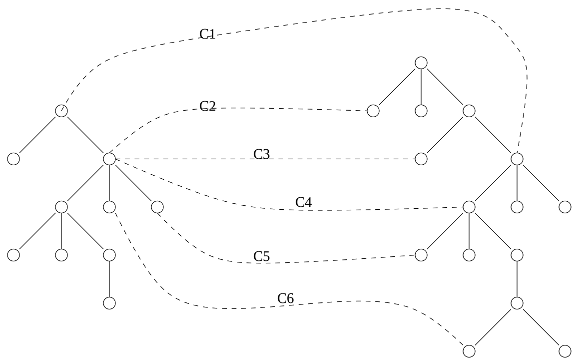

Figure 1: Example of connections between taxonomies.

Constraints are coded with three characters XYZ, which are read as follows: The last character, Z, indicates whether the constraints requires the existence of a connected hypernym (E), hyponym (O), or both (B). The two first characters indicate how the hyper/hyponym relationship is considered in the Spanish taxonomy (character X) and in WordNet (character Y): (I) indicates that only *immediate* hyper/hyponym match, and (A) indicates that *any* ancestor/descendant matches.

Thus, we have constraints IIE/IIO which increase the weight for a connection between a Spanish sense and a WordNet synset when there is a connection between their respective hypernyms/hyponyms. Constraint IIB requires the *simultaneous* satisfaction of IIE and IIO.

Similarly, we have constraints IAE/IAO, which increase the weight for a connection between a Spanish sense and a WordNet synset when there is a connection between the immediate hypernym/hyponym of the Spanish sense and any ancestor/descendant of the WN synset. Constraint IAB requires the *simultaneous* satisfaction of IAE and IAO. Symmetrically, constraints AIE, AIO and AIB, admit recursion on the Spanish taxonomy, but not in WordNet.

Finally, constraints AAE, AAO and AAB, admit recursion on both sides.

For instance, the following example shows a taxonomy in which the IIE constraint would be enough to connect the Spanish node *rapaz* to the <*bird_of_prey*> synset, given that there is a connection between *ave* (hypernym of *rapaz*) and `animal` <*bird*> (hypernym of <*bird_of_prey*>).

animal ⟹(`Tops` <*animal,animate_being,...*>)
⟹(`person` <*beast,brute,...*>)
⟹(`person` <*dunce,blockhead,...*>)
ave ⟹(`animal` <*bird*>)
⟹(`animal` <*fowl,poultry,...*>)
⟹(`artifact` <*bird,shuttle,...*>)
⟹(`food` <*fowl,poultry,...*>)
⟹(`person` <*dame,doll,...*>)
faisan ⟹(`animal` <*pheasant*>)
⟹(`food` <*pheasant*>)
rapaz ⟹(`animal` <*bird_of_prey,...*>)
⟹(`person` <*cub,lad,...*>)
⟹(`person` <*chap,fellow,...*>)
⟹(`person` <*lass,young_girl,...*>)

Constraint IIE would –wrongly– connect the Spanish sense *faisán* to the `food` <*pheasant*> synset, since there is a connection between its immediate hypernym (*ave*) and the immediate hypernym `food` <*pheasant*> (which is `food` <*fowl,poultry,...*>), but the `animal` synsets for *ave* are non–immediate ancestors of the `animal` synsets for <*pheasant*>. This would be rightly solved when using IAE or AAE constraints.

More information on constraints and their application can be found in (Daudé et al., 1999).

4 Experiments and Results

In this section we will describe a set of experiments and the results obtained. A brief description of the used resources is included to set the reader in the test environment.

4.1 Spanish Taxonomies

We tested the relaxation labeling algorithm with the described constraints on a set of disambiguated Spanish taxonomies automatically acquired from monolingual dictionaries. These taxonomies were automatically assigned to a WordNet semantic file (Rigau et al., 1997; Rigau et al., 1998). We tested the performance of the method on two different kinds of taxonomies: those assigned to a well defined and concrete semantic files (`noun.animal`, `noun.food`), and those assigned to more abstract and less structured ones (`noun.cognition` and `noun.communication`).

We performed experiments directly on the taxonomies extracted by (Rigau et al., 1997), as well as on slight variations of them. Namely, we tested on the following *modified* taxonomies:

+top Add a new virtual top as an hypernym of all the top nodes of taxonomies belonging to the same semantic file. The virtual top is connected to the top synset of the WordNet semantic file. In this way, all the taxonomies assigned to a semantic file, are converted to a single one.

no-senses The original taxonomies were built taking into account dictionary entries. Thus, the nodes are not words, but dictionary *senses*. This test consists of collapsing together all the sibling nodes that have the same word, regardless of the dictionary sense they came from. This is done as an attempt to minimize the noise introduced at the sense level by the taxonomy building procedure.

4.2 Bilingual dictionaries

The possible connections between a node in the Spanish taxonomy and WN synsets were extracted from bilingual dictionaries. Each node has as candidate connections all the synsets for all the words that are possible translations for the Spanish word, according to the bilingual dictionary. Although the Spanish taxonomy nodes are dictionary senses, bilingual dictionaries translate words. Thus, this step introduces noise in the form of irrelevant connections, since not all translations necessarily hold for a single dictionary sense.

We used an integration of several bilingual sources available. This multi–source dictionary contains 124,949 translations (between 53,830 English and 41,273 Spanish nouns).

Since not all words in the taxonomy appear in our bilingual dictionaries, coverage will be partial. Table 1 shows the percentage of nodes in each taxonomy that appear in the dictionaries (and thus, that may be connected to WN).

Among the words that appear in the bilingual dictionary, some have only one candidate connection –i.e. are *monosemous*–. Since selecting a connection for these cases is trivial, we will focus on the *polysemic* nodes. Table 2 shows the percentage of polysemic nodes (over the number of words with bilingual connection) in each test taxonomy. The average polysemy ratio (number of candidate connections per Spanish sense) is 15.8, ranging from 9.7 for taxonomies in `noun.animal`, to 20.1 for less structured domains such as `noun.communication`.

4.3 Results

In the performed tests we used simultaneously all constraints with the same recursion pattern. This yields the packs: II*, AI*, IA* and AA*, which were applied to all the taxonomies for the four test semantic files.

Table 3 presents coverage figures for the different test sets, computed as the amount of nodes for which some constraint is applied and thus their weight assignment is changed. Percentage is given over the total amount of nodes with bilingual connections.

To evaluate the precision of the algorithm, we hand checked the results for the *original* taxonomies, using AA* constraints. Precision results can be divided in several cases, depending on the correctness of the Spanish taxonomies used as a starting point.

T_{OK}, F_{OK} The Spanish taxonomy was well built and correctly assigned to the semantic file.

T_{OK}, F_{NOK} The Spanish taxonomy was well built, but wrongly assigned to the semantic file.

T_{NOK} The Spanish taxonomy was wrongly built.

In each case, the algorithm selects a connection for each sense, we will count how many connections are right/wrong in the first and second cases. In the third case the taxonomy was

	original	+top	no-senses
`noun.animal`	45%	45%	43%
`noun.food`	55%	56%	52%
`noun.cognition`	54%	55%	52%
`noun.communication`	66%	66%	64%

Table 1: Percentage of nodes with bilingual connection in each test taxonomy.

	original	+top	no-senses
`noun.animal`	77%	77%	75%
`noun.food`	81%	81%	79%
`noun.cognition`	74%	74%	72%
`noun.communication`	87%	87%	86%

Table 2: Percentage of nodes with more than one candidate connection.

wrongly extracted and is nonsense, so the assignations cannot be evaluated.

Note that we can distinguish right/wrong assignations in the second case because the connections are taken into account over the whole WN, not only on the semantic file being processed. So, the algorithm may end up correctly assigning the words of a hierarchy, even when it was assigned to the wrong semantic file. For instance, in the hierarchy

piel (*skin, fur, peel, pelt*)
$\Longrightarrow$marta (*sable, marten, coal_back*)
$\Longrightarrow$vison (*mink, mink_coat*)

all words may belong either to the semantic file `noun.substance` (senses related to *fur, pelt*) or to `noun.animal` (*animal, animal_part* senses), among others. The right `noun.substance` synsets for each word are selected, since there was no synset for *piel* that was ancestor of the `animal` senses of *marta* and *visón*.

In this case, the hierarchy was well built, and well solved by the algorithm. The only mistake was having assigned it to the `noun.animal` semantic file, so we will count it as a right choice of the relaxation labeling algorithm, but write it in a separate column.

Tables 4 and 5 show the precision rates for each *original* taxonomy. In the former, figures are given over polysemic words (nodes with more than one candidate connection). In the later, figures are computed overall (nodes with at least one candidate connection).

Accuracy is computed at the semantic file level, i.e., if a word is assigned a synset of the right semantic file, it is computed as right, otherwise, as wrong.

To give an idea of the task complexity and the quality of the reported results, even with this simplified evaluation, consider the following:

- Those nodes with only one possible synset for the right semantic file (30% in average, ranging from 22% in `noun.communication` to 45% in `noun.animal`) are not affected by the evaluation at the semantic file level.
- The remaining nodes have more than one possible synset in the right semantic file: 6.3 in average (ranging from 3.0 for `noun.animal` to 8.7 for `noun.communication`).
- Thus, we can consider that we are evaluating a task easier than the actual one (the actual evaluation would be performed at the synset level). This simplified task has an average polysemy of 6.7 possible choices per sense, while the actual task at the synset level would have 15.8. Although this situates the baseline of a random assignment about 15% instead of 6%, it is still a hard task.

5 Conclusions

We have applied the relaxation labeling algorithm to assign an appropriate WN synset to each node of an automatically extracted taxonomy. Results for two different kinds of conceptual structures have been reported, and they point that this may be an accurate and robust method (not based on ad-hoc heuristics) to connect hierarchies (even in different languages).

WN file	taxonomy	II*	AI*	IA*	AA*
noun.animal	original	134 (23%)	135 (23%)	357 (62%)	365 (63%)
	+top	138 (24%)	143 (25%)	375 (65%)	454 (78%)
	no-senses	118 (23%)	119 (20%)	311 (61%)	319 (62%)
noun.food	original	119 (36%)	130 (39%)	164 (49%)	180 (63%)
	+top	134 (40%)	158 (47%)	194 (58%)	259 (77%)
	no-senses	102 (36%)	111 (39%)	153 (51%)	156 (55%)
noun.cognition	original	225 (37%)	230 (38%)	360 (60%)	373 (62%)
	+top	230 (38%)	240 (40%)	395 (65%)	509 (84%)
	no-senses	192 (37%)	197 (38%)	306 (59%)	318 (61%)
noun.communication	original	552 (43%)	577 (45%)	737 (57%)	760 (59%)
	+top	589 (46%)	697 (54%)	802 (62%)	1136 (88%)
	no-senses	485 (43%)	509 (45%)	645 (57%)	668 (59%)

Table 3: Coverage of each constraint set for different test sets.

	precision over T_{OK}, F_{OK}	precision over T_{OK}, F_{NOK}	total precision over T_{OK}	number of T_{NOK}
animal	279 (90%)	30 (91%)	309 (90%)	23
food	166 (94%)	3 (100%)	169 (94%)	2
cognition	198 (67%)	27 (90%)	225 (69%)	49
communication	533 (77%)	40 (97%)	573 (78%)	16

Table 4: Precision results over polysemic words for the test taxonomies.

The experiments performed up to now seem to indicate that:

- The relaxation labeling algorithm is a good technique to link two different hierarchies. For each node with several possible connections, the candidate that best matches the surrounding structure is selected.
- The only information used by the algorithm are the hyper/hyponymy relationships in both taxonomies. These local constraints are propagated throughout the hierarchies to produce a global solution.
- There is a certain amount of noise in the different phases of the process. First, the taxonomies were automatically acquired and assigned to semantic files. Second, the bilingual dictionary translates words, not senses, which introduces irrelevant candidate connections.
- The size and coverage of the bilingual dictionaries used to establish the candidate connections is an important issue. A dictionary with larger coverage increases the amount of nodes with candidate connections and thus the algorithm coverage

6 Proposals for Further Work

Some issues to be addressed to improve the algorithm performance are the following:

- Further test and evaluate the precision of the algorithm. In this direction we plan –apart from performing wider hand checking of the results, both to file and synset level– to use the presented technique to link WN1.5 with WN1.6. Since there is already a mapping between both versions, the experiment would provide an idea of the accuracy of the technique and of its applicability to different hierarchies of the same language. In addition, it would constitute an easy way to update existing lexical resources.
- Use other relationships apart from hyper/hyponymy to build constraints to select the best connection (e.g. sibling, cousin, synonymy, meronymy, etc.).
- To palliate the low coverage of the bilingual dictionaries, candidate translations could be inferred from connections of surrounding senses. For instance, if a sense has no candidate connections, but its hypernym

	precision over T_{OK}, F_{OK}	precision over T_{OK}, F_{NOK}	total precision over T_{OK}
`animal`	424 (93%)	62 (95%)	486 (93%)
`food`	166 (94%)	83 (100%)	149 (96%)
`cognition`	200 (67%)	245 (99%)	445 (82%)
`communication`	536 (77%)	234 (99%)	760 (81%)

Table 5: Precision results over all words for the test taxonomies.

does, we could consider as candidate connections for that node all the hyponyms of the synset connected to its hypernym.

- Use the algorithm to enrich the Spanish part of EuroWordNet taxonomy. It could also be applied to include taxonomies for other languages not currently in the EWN project.

In addition, some ideas to further exploit the possibilities of these techniques are:

- Use EWN instead of WN as the target taxonomy. This would largely increase the coverage, since the candidate connections missing in the bilingual dictionaries could be obtained from the Spanish part of EWN, and viceversa. In addition, it would be useful to detect *gaps* in the Spanish part of EWN, since a EWN synset with no Spanish words in EWN, could be assigned one via the connections obtained from the bilingual dictionaries.
- Since we are connecting dictionary senses (the entries in the MRD used to build the taxonomies) to EWN synsets: First of all, we could use this to disambiguate the right sense for the genus of an entry. For instance, in the Spanish taxonomies, the genus for the entry *queso_1* (cheese) is *masa* (mass) but this word has several dictionary entries. Connecting the taxonomy to EWN, we would be able to find out which is the appropriate sense for *masa*, and thus, which is the right genus sense for *queso_1*. Secondly, once we had each dictionary sense connected to a EWN synset, we could enrich EWN with the definitions in the MRD, using them as Spanish glosses.
- Map the Spanish part of EWN to WN1.6. This could be done either directly, or via mapping WN1.5–WN1.6.

7 Acknowledgments

This research has been partially funded by the Spanish Research Department (ITEM Project TIC96-1243-C03-03), the Catalan Research Department (CREL project), and the UE Commission (EuroWordNet LE4003).

References

A. Ageno, I. Castellón, F. Ribas, G. Rigau, H. Rodríguez, and A. Samiotou. 1994. TGE: Tlink Generation Environment. In *Proceedings of the 15th International Conference on Computational Linguistics (COLING'94)*, Kyoto, Japan.

M. Alvar, editor. 1987. *Diccionario General Ilustrado de la Lengua Española VOX*. Biblograf S.A, Barcelona, Spain.

J. Atserias, S. Climent, X. Farreres, G. Rigau, and H. Rodríguez. 1997. Combining Multiple Methods for the Automatic Construction of Multilingual WordNets. In *proceedings of International Conference on Recent Advances in Natural Language Processing (RANLP'97)*, Tzigov Chark, Bulgaria.

J. Daudé, L. Padró, and G. Rigau. 1999. Experiments on Applying Relaxation Labeling to Map Multilingual Hierarchies. Technical Report LSI-99-5-R, Departament de LSI. Universitat Politècnica de Catalunya.

J. O. Eklundh and A. Rosenfeld. 1978. Convergence Properties of Relaxation Labelling. Technical Report 701, Computer Science Center. University of Maryland.

X. Farreres, G. Rigau, and H. Rodríguez. 1998. Using WordNet for Building WordNets. In *Proceedings of COLING-ACL Workshop on Usage of WordNet in Natural Language Processing Systems*, Montréal, Canada.

K. Knight and S. Luk. 1994. Building a Large-Scale Knowledge Base for Machine Translation. In *Proceedings of the American Association for Artificial Inteligence (AAAI'94)*.

L. Màrquez and L. Padró. 1997. A Flexible POS Tagger Using an Automatically Acquired Language Model. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics. Joint ACL/EACL*, pages 238–245, Madrid, Spain, July.

G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller. 1991. Five Papers on WordNet. *International Journal of Lexicography.*

A. Okumura and E. Hovy. 1994. Building japanese-english dictionary based on ontology for machine translation. In *proceedings of ARPA Workshop on Human Language Technology*, pages 236–241.

L. Padró. 1998. *A Hybrid Environment for Syntax–Semantic Tagging.* Phd. Thesis, Dep. Llenguatges i Sistemes Informàtics. Universitat Politècnica de Catalunya, February. http://www.lsi.upc.es/˜padro.

P. Procter, editor. 1987. *Longman Dictionary of Common English.* Longman Group, Harlow, Essex, England.

J. Richards, D. Landgrebe, and P. Swain. 1981. On the accuracy of pixel relaxation labelling. *IEEE Transactions on Systems, Man and Cybernetics*, 11(4):303–309.

G. Rigau, H. Rodríguez, and J. Turmo. 1995. Automatically extracting Translation Links using a wide coverage semantic taxonomy. In *proceedings 15th International Conference AI'95*, Montpellier, France.

G. Rigau, J. Atserias, and E. Agirre. 1997. Combining Unsupervised Lexical Knowledge Methods for Word Sense Disambiguation. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics. Joint ACL/EACL*, pages 48–55, Madrid, Spain, July.

G. Rigau, H. Rodríguez, and E. Agirre. 1998. Building Accurate Semantic Taxonomies from MRDs. In *Proceedings of COLING-ACL'98*, Montréal, Canada.

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

M. Utiyama and K. Hasida. 1997. Bottom-up Alignment of Ontologies. In *Proceedings of IJCAI workshop on Ontologies and Multilingual NLP*, Nagoya, Japan.

A. Voutilainen and L. Padró. 1997. Developing a Hybrid NP Parser. In *Proceedings of the 5th Conference on Applied Natural Language Processing, ANLP*, pages 80–87, Washington DC. ACL.