PHORA: A system to solve the Anaphora in Spanish

M. Palomar, M. Saiz-Noeda, R. Muñoz, A. Suárez and P. Martínez-Barco
Dpto. Lenguajes y Sistemas Informáticos, Universidad de Alicante, Spain
{mpalomar,max,rafael,armando,patricio}@dlsi.ua.es

Abstract

In this paper we present a whole Natural Language Processing (NLP) system for Spanish. The core of this system is the parser, which uses the grammatical formalism Lexical-Functional Grammars (LFG). Another important component of this system is the anaphora resolution module. To solve the anaphora, this module contains a method based on linguistic information (lexical, morphological, syntactic and semantic), structural information (anaphoric accessibility space in which the anaphor obtains the antecedent) and statistical information. This method is based on constraints and preferences and solves pronouns and definite descriptions. Moreover, this system fits dialogue and non-dialogue discourse features. The anaphora resolution module uses several resources, such as a lexical database (Spanish WordNet) to provide semantic information and a POS tagger providing the part of speech for each word and its root to make this resolution process easier.

Keywords: Anaphora resolution, semantics, LFG grammar and parsing, EuroWordNet.

1. Introduction

It is necessary to have adequate information sources in order to develope a suitable mechanism for anaphora resolution. In the course of the last years, numerous researches have concentrated their efforts on solving the problem of lexical and morphological analysis. Also, they have been worked on the obtaining of universal information sources to provide the information adapted to each problem. With reference to the addition of semantic information in the resolution of linguistic phenomena, it is not easy to find references and resources that provide succesfull results.

The system presented in this work combines a set of tools (taggers, parser) and uses a combination of information sources to anaphora resolution (linguistic and structural information).

Summarizing, we propose a NLP system that counts on as much information sources as possible for the anaphora resolution. We think that the base of the language processing is the information used for it.

Next section shows the main aspects of the anaphora phenomenon. Following, the paper gives a detailed description of the NLP system and all the processes and resources related to it. After that, the constraint-based mechanism for the anaphora resolution and the required information sources are explained. Finally, some conclusions of this work and the work in progress are presented.

2. The problem

We commonly use different expressions to refer to a person, an object, an event, a place or a process. These expressions usually used in this way are, among others, pronouns and definite descriptions[1]. But, the latter not always make reference to a linguistic entity previously mentioned. In this case, the definite description introduces a new entity into the discourse.

This new entity (person, object, event, place or event) is called *candidate* of the next anaphoric expressions. The candidates are noun phrases, verbal phrases, full sentences or text fragments. According to Eckert and Strube (1999), if the antecedent is a noun phrase then the anaphora is classified as *individual anaphora*, otherwise, the anaphora is classified as *abstract anaphora*.

Moreover, different kinds of relations can be found between the antecedent and the anaphoric expression (pronoun and definite description). These relations can be *parts of*, *set-subset* and *set-member* for definite descriptions and *identity* for both pronouns and definite descriptions. Another difference between pronouns and definite descriptions is the *accessibility space* (see section 4.5) used to solve the references. Pronouns use a shorter accessibility space than definite description because the former don't provide semantic information.

In this paper we only treat identity relations between the antecedent and the anaphoric expression, particularly those whose antecedent is a noun phrase (individual anaphora).

Pronouns and definite descriptions can be classified into different types. Pronouns are divided into personal, demonstrative, reflexive and omitted pronouns (zero-pronouns) in third person in Spanish. Personal and demonstrative pronouns are classified according to whether they are included in a prepositional phrase (PP) or not and whether they are complement personal pronouns. Moreover, the antecedents (noun phrases) can be divided into intrasentential and intersentential. Definite descriptions are divided into non-anaphoric and anaphoric definite descriptions. The former introduce a new entity discourse while the latter refers to a previous antecedent. A complete classification of definite description has been developed by several authors, such as Hawkins (1978), Christopherson (1939) for English language and Muñoz et *al.* (2000) for Spanish language.

The following examples give a short view of the different types of anaphora above mentioned:

- *Complement personal pronoun*. Ana abre <u>la puerta</u> y <u>la</u> cierra tras de sí.
- *Personal pronouns not included in a PP*. <u>Andrés</u> es mi vecino. <u>Él</u> vive en el segundo.

[1]Rusell (Russell, 1919) called definite descriptions to definite noun phrases. That is, noun phrases headed by definite article or demonstratives, such as *the newspaper* or *this journal*

- *Personal pronouns included in a PP*. Juan debe asistir pero Pedro lo hará por él
- *Demonstrative Pronoun not included in a PP*. Pedro está enfadado con Antonio. Éste no le habla.
- *Demonstrative Pronoun included in a PP*. Ana vive con Paco y cocina para él cada día.
- *Reflexive Pronoun*. Ana abre la puerta y la cierra tras de sí
- *Omitted pronoun (zero pronouns*. Ana abre la puerta y Ø la cierra tras de sí
- *Definite description with the same head its antecedent*. La casa de la playa es preciosa ... Era la casa de un mdico
- *Bridging reference*[2]. Los soldados irrumpieron en la ciudad. Esos bravos hombres mataron a los luchadores de la guerrilla.

Next section shows the different resources used by the PHORA system.

3. Resources

The system uses the following resources:

- **EuroWordNet.** WordNet (WN), as described by Miller et *al.* (1993), is an electronic dictionary that stores sets of synonyms. Each set of synonyms - synset - describes a semantic concept and contains a list of pairs (word - sense number) and several pointers to others synsets as semantic relations. So, meanings of a word are stored in WordNet as synsets, one per sense. Furthermore, each synset can have a definition or gloss, like conventional dictionaries.

  EuroWordNet (EWN) (Vossen, 1998) is a recent development based on the English WordNet 1.5. Briefly, EWN consists of a set of different WordNets for various languages (English, Dutch, Spanish, Italian, German, French, Czech and Estonian) and an inter-lingual index that links the synsets of each language with the synsets in the English WN. In most cases, these synsets have the same meaning. Recently, WordNets in other languages have been developed. Like WN1.5, each WN in EWN maintains a set of pointers between synsets as a representation of heterogeneous semantic relations like hypernym/hyponym between nouns. Moreover, the hyper/hyponimy relation establishes a basic ontology (Top Ontology) common to all the languages, classifying the synsets in a conceptual category.

  EuroWorNet will be used as a base resource for semantic tagging in order to apply semantic knowledge to the anaphora resolution process through semantic patterns and semantic constraints an preferences.

- **Semantic tagger** In order to help many NLP tasks, particularly anaphora resolution, a semantic tagger is being developed. This module follows the last proposals and developments in word sense dissambiguation (WSD) and makes use of various and heterogeneous sources of information (morphological, syntactic and grammatical information, selectional preferences, constraints, etc.). Particularly, this module will tag nouns and verbs with EWN Top Ontology classes and synset level tags. Many WSD algorithms are being evaluated at this moment in order to find the best option, methods like Relaxation Labelling (Padró, 1998), Maximum Entropy (Ratnaparkhi, 1998), Memory Based algorithms (Veenstra et al., 2000), and Boosting (Escudero et al., 2000). In general, these are supervised learning methods, which make use of Machine Learning algorithms and electronic dictionaries like EWN.

- **Parser.** The parser takes different input sources. On the one hand, it receives the output of a POS tagger that provides, for each word in the analysed corpus, its morphological features (gender, number, person, verbal tense,...). On the other hand, the parser takes as input the output of the semantic tagger. The parser uses the grammatical formalism LFG for the analysis process.

- **LFG Grammar.** Dalrymple *et al* (1997) say that LFG assumes two syntactic levels of representation. Constituent structure (c-structure) encodes phrasal dominance and precedence relations, and is represented as a phrase structure tree. Functional structure (f-structure) encodes syntactic predicate-argument structure, and is represented as an attribute-value matrix.

  F-structure consists of a collection of attributes, such as PRED, SUBJ, OBJ or IOBJ, whose values can be other f-structures. In this paper, we present the structures suitable for pronouns and definite descriptions, which are the required components for the anaphora resolution.

  Figure 1 shows the output of the parser using the gramatical formalism LFG.

- **The anaphoric accessibilitty space.** NLP researchers have proposed a lot of mechanisms to solve the anaphora problem focusing their interest in the choice of the adequate antecedent among a list of candidates. However, there is a lack of studies performing a proposal about anaphoric accessibility space, that is the space where the system can found this list of candidates in each kind of anaphora. This definition has a great importance in the remaining process because definitions of anaphoric accessibility space which are too short cause the removal of valid antecedents for the anaphor. On the other hand, definitions of anaphoric accessibility space which are too large cause large candidate lists, where failure probabilities in anaphora resolution increase.

  Anaphora resolution systems based on linguistic knowledge usually define an accessibility space using

[2]Clark (1977) called bridging descriptions to definite descriptions that either have an antecedent denoting the same discourse entity, but using a different head noun (synonym, hypernym or hyponym) or are related by other relation than identity

```
Sentence: "Portugal ha vivido en torno a su equipo"

S[ aTree : "[[ProperNoun]NP
        [AuxVerb MainVerb]VC
        [Preposition
            [[[PossessiveAdj]Det]
             Determine Noun]NP
        ]PP
]S",

adjunct : <:
    PP[
        head : Noun[
            exp : "equipo",
            lex : LexMorph,
            infl : InflMorph[
                number : #143=Singular,
                gender : #141=Masculine]],
        mod : ADJP,
        prep : Preposition[
            exp : "en_torno_a"],
        spec : Determine[
            number : #143,
            gender : #141,
            det : Det[
                number : #143,
                gender : #141,
                head : PossessiveAdj[
                    exp : "su",
                    lex : LexMorph,
                    infl : InflMorph[
                        number : #143,
                        gender : #141]]],
            postdet : Null,
            predet : Null]]:>,
head : VC[
        tense : #293=Present,
        number : #283=Singular,
        mod1 : AuxVerb[
            exp : "haber",
            infl : InflMorph[
                tense : Present,
                person : Third,
                number : Singular,
                mood : Indicative]],
        head : MainVerb[
            exp : "vivir",
            lex : LexMorph[
                subcat : #252=Transitive],
            infl : InflMorph[
                tense : #293,
                number : #283,
                mood : Participle]],
        cat : #252],
subject : NP[
        number : #283,
        head : ProperNoun[
            exp : "portugal"]]]
```

Figure 1: *Parser output*

n previous sentences to the anaphor, where *n* is variable according to the kind of anaphora. However, this definition is arbitrary and there are not structural reasons for it.

In Martínez-Barco and Palomar (2000a), a study about the anaphoric accessibility space is presented. This study shows that antecedents of pronominal and adjectival anaphors in Spanish can almost always be found in the set of noun phrases taken from the anaphoric accessibility space, defining this space according to an structure based on adjacency pairs and topics.

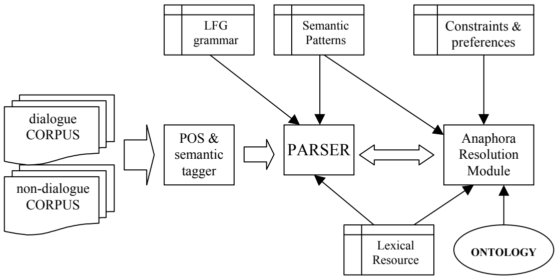

Figure 2: *PHORA System Architecture*

According to Fox (1987) the first mention of a referent in a sequence is done with a full noun phrase. After that, by using an anaphor the speaker displays an understanding that sequence has not been closed down. The author shows that two different sequences generate most of the anaphors to be found in dialogues: the adjacency pair and the topic scope. The former generates references to any local noun phrase, and the latter generates references to the main topic of the dialogue.

Based on this, the anaphoric accessibility space is proposed as the set of noun phrases taken from:

- the same adjacency pair as the anaphor, and
- the previous adjacency pair to the anaphor, and
- another adjacency pair including the anaphor adjacency pair, and
- the noun phrase representing the main topic of the dialogue.

According to this study, 95.9% of the antecedents were located in the proposed anaphoric accessibility space. Remaining antecedents (4.1%) were estimated to be located in subtopics of the dialogues.

Besides, this study shows the importance of defining the adequate anaphoric accessibility space for each kind of anaphora, and also for each kind of discourse.

4. PHORA System

PHORA system is a whole Natural Language Processing system made up by three main modules as shown in figure 2. The core of this system is the parser which uses the grammatical formalism Lexical-Functional Grammars (LFG). Another important component of this system is the anaphora resolution module. To solve the anaphora, this module contains a method based on linguistic information (lexical, morphological, syntactic and semantic), structural information (anaphoric accessibility space in which the anaphor obtains the antecedent) and statistical information. This method is based on constraints and preferences and solves pronouns and definite descriptions. Moreover, this system fits dialogue and non-dialogue discourse features.

4.1. Anaphora Resolution Module

Our method of anaphora resolution within this module, is based on constraints and preferences that are extracted from linguistic and structural sources. The method solves the anaphora phenomena both in dialogue and non-dialogue discourses.

The method uses an anaphoric accessibility space in order to extract the list of possible candidates to be the solution of the anaphor. This anaphoric accessibility space varies depending on the kind of discourse (dialogue or non-dialogue) and the kind of anaphora (pronoun or definite description). Anaphora resolution in dialogues uses an anaphoric accessibility space based on the dialogue structure. However, the anaphoric accessibility space used in non-dialogue discourse is based on a window of sentences when it is applied to pronouns and is based on the previous full text when applied to definite descriptions (Martínez-Barco and Palomar, 2000b).

Next, we will describe our sets of constraints and preference that are defined for each kind of discourse and for each kind of anaphor. Finally, the system that manages these sets will be shown.

The sets of constraints and preferences are based on different kinds of knowledge, excepting the discourse structure-based knowledge that requires not only semantic knowledge but also word knowledge and an almost perfect parsing (Azzam et al., 1998).

Like in the known systems that are based in constraints and preferences, the former eliminate all the candidates that disagree with the anaphor, while the latter treats to establish the preference order among the candidates that have been "survived" to the constraints. Next subsections describe the constraints and the preferences used in the system.

4.1.1. Constraints

- ***Morphological constraints:*** Morphological constraints establish gender, number and person parallelisms that demand the compatibility or agreement between the antecedent and the anaphoric pronoun. If the anaphoric expression is a definite description this constrain is not applied because the correct antecedent and the definite description can have different morphological structure.

- ***Syntactic constraints:*** Syntactic constraints are based on the well-known non-coreference conditions by Lappin and Leass (1994). We propose conditions for NP-pronoun non-coreference adapted for Spanish. A pronoun is non-coreferential with a noun phrase (NP) if any of the non-coreference conditions are fulfilled. These conditions relate reflexive, demonstrative and personal pronouns with their syntactic role and position in the sentence, as can be seen in Palomar *et al* (2000).

- ***Semantic constraints:***

  Semantic knowledge is applied for the anaphora resolution method in two different ways. On the one hand, for pronoun resolution, from the semantic features associated to each antecedent NP, semantic constraints eliminate those candidates that are not compatible with the verb of the sentence in which the anaphoric expression appears. So, in this case, the compatibility is not determined directly by the anaphoric expression, but by the verb that it accompanies. Subsection 4.2. makes a further detailed explanation of the pattern learning method that is used to determine the verb compatibility.

  On the other hand, semantic relations such as synonym, hypernym, hyponym and meronym are applied between both antecedent and definite description in order to provide the candidates.

4.1.2. Preferences

According to the kind of discourse and the kind of anaphor a subset of preferences will be applied.

- **Repetition.** Candidates with the same head noun as their definite description and with the same pre and post-modifiers. This preference is used to solve references produced by definite descriptions. The preferences used to solve definite description can be seen in Muñoz et *al.* (2000) and Muñoz and Palomar (2000).

- **Same Head.** Candidates with the same head noun as their definite description and with pre and post-modifiers semantically related. This preference is used to solve references produced by definite descriptions.

- **Related Semantically Head.** Candidates with a head noun semantically related with the head noun of definite description. This preference is used to solve references produced by definite descriptions.

- **Role Agent.** Candidates that are related with the verb of the sentence in which the definite description appears. If this relationship is direct, then candidates with a subject syntactic function are preferred. Otherwise, if this relationship is indirect, then those with a complement function are preferred. This preference is used to solve references produced by definite descriptions.

- **Same Adjacency Pair.** Candidates that are in the same AP as the anaphor. This preference is used to solve references produced by pronouns in dialogues.

- **Previous Adjacency Pair.** Candidates that are in the previous AP to the anaphor. This preference is used to solve references produced by pronouns in dialogues.

- **Nested Adjacency Pair.** Those candidates that are in the most recent unclosed AP. This preference is used to solve references produced by pronouns in dialogues.

- **Same Position with reference to the verb.** Those candidates that are in the same position with reference to the verb as the anaphor (before or after). This preference is used to solve references produced by pronouns.

- **Same Syntactic Position.** Those candidates in the same syntactic position as the anaphor. This preference is used to solve references produced by pronouns.

- **Closest.** Nearest candidate. This preference is used by any kind of anaphoric expression and discourse. Moreover, this preference guarantees that only one candidate will be selected at the end of the process. This candidate will be the antecedent proposed.

4.2. Pattern Learning method

For defining the semantic structure of a sentence, we have considered a set of patterns formed by the semantic concepts associated to the main elements in the aforesaid sentence. This way, the objective of this method is to extract a set of *subject noun-verb* and *verb-complement noun* patterns from each sentence. The patterns will be formed by the semantic or ontological concepts of the noun phrase head (subject or complement) and predicate verbal phrase head. For the pattern construction, the Spanish version of the lexical resource WordNet, within the EuroWordNet proyect described by Vossen (1998) has been chosen. WordNet provides a main level of ontological concepts to describe all the words contained in the knowledge base. These concepts are 25 for nouns and 15 for verbs and they get the main semantic characteristic of each word sense. These concepts are:

- *Nouns*: act, animal, artifact, attribute, body, cognition, communication, event, feeling, food, group, location, motive, object, person, phenomenon, plant, possession, process, quantity, relation, shape, state, substance and time

- *Verbs*: body, change, cognition, communication, competition, consumption, contact, creation, emotion, motion, perception, possession, social, stative and weather

For the pronoun resolution system, it is important to define the concept of compatibility between a noun (subject or complement) and a verb. This compatibility will allow choosing the correct antecedent of an anaphora among a group of noun phrases, being the correct antecedent the one with the highest compatibility degree. For example, taking the verb (comer) (*to eat*) and two noun phrases la piedra (*the stone*) and el león (*the lion*), the patterns generated by both nouns and the verb are *object-consumption* and *animal-consumption*. Intuitively, we can deduce that the second pattern defines elements as semantically more compatible. So, if both noun phrases comprise an antecedent list of an anaphoric expression with the verb comer, it is possible to say that, from the semantic point of view, el león can be the correct antecedent because of its compatibility.

4.3. Anaphora resolution process

The anaphora resolution module uses as input the output of the parser. This output is made up by different kind of information (morphological, syntactic and semantic) for each word and for each group of word (chunks). The anaphora resolution module uses this information to apply a set of constraints and preferences. Depending on the kind of discourse (dialogue or non-dialogue) and the kind of anaphoric expression (pronoun or definite description), the module uses a different set of constraints and preferences. But, the method to be applied is the same without bearing this difference in mind.

Once the morphological, syntactical and semantic constraints have been applied, if there is only one candidate left, it is chosen as the correct antecedent of the pronoun.

Otherwise, a set of preferences are applied to provide the antecedent of anaphora. In this line, we can remark the use of a set of semantic criteria based on the semantic patterns and the semantic structure above mentioned is applied.

The use of these patterns is double. We can use this semantic information as a constraint and also as a preference. It depends on the compatibility degree between the anaphora and the antecedent. It is considered that a compatibility degree of 0% can be used as a constraint. In case of a lower compatibility degree the candidates will "compete" with their compatibility degree for being the correct antecedent.

This semantic information will allow the selection of the antecedent that is conceptually most compatible. Some results concerning to the use of these patterns in Spanish and English texts can be seen in Saiz-Noeda et *al.* (1999; 2000) respectively, while another apporach to apply semantic information from an ontology instead of WordNet can be found in Saiz-Noeda and Palomar (2000).

5. Conclusions

General-purpose Natural Language Processing (NLP) systems that can be applied to any domain are nowadays more and more sought after. Also, these systems need to use heterogeneous information sources to achieve optimal results. It means, the use of different kind of information is necessary in the treatment of Natural Language.

In this work, we have presented a general Natural Language Processing system with a module to solve a discourse phenomenon (anaphora) that can be applied to different applications, such as, Machine Translation (MT), Information Retrieval (IR) or Information Extraction (IE).

We have developed this anaphora resolution module based on constraints and preferences and it can be applied to solve the anaphora in dialogue and non-dialogue texts. Moreover, the system can solve different types of anaphoric expressions such as pronouns and definite descriptions.

Last years, many research efforts have been focused on anaphora resolution in this kind of expressions achieving satisfactory results for Spanish. Nowadays, we are developing tools to link all information sources in an only independent multi-platform system[3] to be applied in any application.

6. References

S. Azzam, K. Humphrey, and R. Gaizauskas. 1998. Evaluating a Focus-Based Approach to Anaphora Resolution. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING-ACL'98)*, pages 74–78, Montreal (Canada), August.

[3]More information about these tools can be seen at http://gplsi.dlsi.ua.es

Dimitris N. Christodoulakis, editor. 2000. *Second International Conference on Natural Language Processing, NLP'2000*, volume 1835 of *Lectures Notes in Artificial Intelligence*, Patras, Greece, June. Springer-Verlag.

P. Christopherson. 1939. *The Articles: A study of their theory and use in English*. E. Munksgaard, Copenhagen.

H. H. Clark. 1977. Bridging. In P. Johnson-Laird and P Wason, editors, *Thinking: readings in cognitive science*, pages 411–420. Cambridge University Press, London New York.

M. Dalrymple, J. Lamping, F. Pereira, and V. Saraswat. 1997. Quantifiers, Anaphora, and Intensionality. *Journal of Logic, Language, and Information*, (6):219–273.

M. Eckert and M. Strube. 1999. Resolving Discourse Deitic Anaphora in Dialogues. In *Proceedings of 9th Conference of the European Chapter of the Association for Computational Linguistics (EACL'99)*, pages 37–44, Bergen, Norway.

G. Escudero, L. Márquez, and G. Rigau. 2000. Boosting Applied to Word Sense Disambiguation. In R. L. de Mntaras and E. Plaza, editors, *Proceedings of the 11th European Conference on Machine Learning, ECML 2000*, Lecture Notes in Artificial Intelligence, Barcelona, Spain, September. Springer Verlag.

B. Fox. 1987. *Discourse Structure and Anaphora.* Written and conversational English. Cambridge Studies in Linguistics. Cambridge University Press, Cambridge.

J. A. Hawkins. 1978. *Definiteness and indefiniteness*. Humanities Press, Atlantic Highlands, New Jersey.

S. Lappin and H. Leass. 1994. An algorithm for pronominal anaphora resolution. *Computational Linguistics*, 20(4):535–561.

P. Martínez-Barco and M. Palomar. 2000a. Empirical study of the anaphoric accessibility space in Spanish dialogues. In *Proceedings of the Fourth Workshop on the Semantics and Pragmatics of Dialogue (GOTALOG'2000)*, Gteborg, Sweden, June.

P. Martínez-Barco and M. Palomar. 2000b. Structure-based Anaphoric Accesibility Space. *Computational Linguistics*. Submitted.

G. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller. 1993. Five Papers on WordNet. *International journal of lexicography*, Special Issue(3).

R. Muñoz and A. Ferrández. 2000. Definite Descriptions Resolution in Spanish. In *Proceeding of ACIDCA'2000: Corpora and Natural Language Processing*, pages 140–145, Monastir, Túnez, Marzo.

R. Muñoz and M. Palomar. 2000. Processing of Spanish Definite Descriptions with the Same Head. In Christodoulakis (Christodoulakis, 2000), pages 212–220.

R. Muñoz, M. Palomar, and A. Ferrández. 2000. Processing of Spanish Definite Descriptions. In O. Cairo and E. L. Sucar and F. J. Cantu, editor, *Proceeding of Mexican International Conference on Artificial Intelligence*, volume 1793 of *Lectures Notes in Artificial Intelligence*, pages 526–537, Acapulco, Mexico, April. Springer-Verlag.

L. Padró. 1998. *A Hybrid Environment for Syntax Semantic Tagging*. Ph.D. thesis, Departament Llenguatges i Sistemes Informatics. Universitat Politecnica de Catalunya, Barcelona.

M. Palomar, A. Ferrández, L. Moreno, P. Martínez-Barco, J. Peral, M. Saiz-Noeda, and R. Muñoz. 2000. An algorithm for Anaphora Resolution in Spanish Texts. *Computational Linguistics (submitted)*.

A. Ratnaparkhi. 1998. *Maximum Entropy Models for Natural Language Ambiguity Resolution*. Ph.D. thesis, University of Pennsylvania.

B. Russell. 1919. Descripciones. In Luis Ml. Valdés Villanueva, editor, *La búsqueda del significado: Lecturas de filosofía del lenguaje*, pages 46–56. Tecnos.

M. Saiz-Noeda and M. Palomar. 2000. Semantic Knowledge-driven Method to Solve Pronominal Anaphora in Spanish. In Christodoulakis (Christodoulakis, 2000), pages 204–211.

M. Saiz-Noeda, A. Surez, and J. Peral. 1999. Propuesta de incorporacin de informacin semntica desde WordNet al anlisis sintctico parcial orientado a la resolucin de la anfora. *Procesamiento del Lenguaje Natural*, 25:167–173.

M. Saiz-Noeda, J. Peral, and A. Surez. 2000. Semantic Compatibility Techniques for Anaphora Resolution. In *Proceeding of ACIDCA'2000: Corpora and Natural Language Processing*, pages 43–48, Monastir, Túnez, Marzo.

J. Veenstra, A. van den Bosch, S. Buchholz, W. Daelemans, and J. Zavrel. 2000. Memory-Based Word Sense Disambiguation. *Computers and the Humanities*, 34(1-2):171–177.

P. Vossen. 1998. EuroWordNet: Building a Multilingual Database with WordNets for European Languages. *The ELRA Newwsletter. K.Choukri, D.Fry, M. Nilsson (des). ISSN:1026-8200*, 3(1).