Proceedings of VEXTAL ’99 – Venezia, San Servolo, V.I.U. – 22-24 November 1999

Evaluation of pronoun resolution algorithm for Spanish dialogues

Patricio Martínez-Barco[†], **Rafael Muñoz**[†], **Saliha Azzam**[‡], **Manuel Palomar**[†] and **Antonio Ferrández**[†]

[†]Research Group on Language Processing and Information Systems.
Department of Software and Computing Systems. University of Alicante
03690 San Vicente del Raspeig. Alicante, Spain,
Phone: +34 96 590 34 00 Fax: +34 96 590 93 26
{ patricio, mpalomar, antonio } @dlsi.ua.es
[‡]Dept. of Computer Science. University of Sheffield
Regent Court, Portobello Rd Sheffield S1 4DP UK
Phone: +44 114 222 1909 Fax: +44 114 222 1810
saliha@dcs.shef.ac.uk

Abstract

In this paper, we present an algorithm for anaphora resolution in Spanish dialogues and an evaluation of the algorithm for pronominal anaphora. The proposed algorithm uses both linguistic information and the structure of the dialogue to find the antecedent of the anaphors. The system has been evaluated on ten dialogues.

1. Introduction

In this paper, we present an algorithm for anaphora resolution in Spanish dialogues and an evaluation of the algorithm for pronominal anaphora. The presented work represents one first work done in the study of pronominal anaphora resolution in dialogues for the Spanish language. The proposed algorithm combines two kinds of information. On one hand, it uses linguistic information to accept or reject the antecedent of the anaphors. On the other hand, it uses the structure of the discourse. We consider that dialogues, due to their structure and unlike non-dialogue discourses, have different spaces of anaphoric accessibility.

Using the structure of the discourse, and its segmentation for the analysis of anaphoric relations, has been widely exploited in the literature, see for example by Webber (1988) that analyses reference to discourse segments in a text, and more generally Grosz et *al.* (1995) that defined the centering theory based on the discourse structure study proposed earlier in Grosz and Sidner (1986). This work had led to several other studies exploring the centering theory see for example Manabu and Kouji (1996) and Walker (1997), or Azzam et *al.* (1998) that applied a focus-based approach (i.e, using the notion of *center* or *focus*) for pronoun resolution, and more particulary Byron and Stent (1998) that examines some issues of using the centering theory for dialogues.

For dialogues, based on Grosz and Sidner (1986) definitions, Rocha (1998) justifies the need to carry out a segmentation of the dialogue where each segment is associated with a certain set of focus elements, in order to study anaphoric relations correctly. Rocha (1998) develops an annotation scheme according to the topic[1] structure of the dialogue. Firstly, Rocha establishes a global topic for each fragment of his corpus. Next, he divides the text into segments according to a local topic: when the local topic changes, a new segment is established. Segment and topic changes are manually annotated on the corpus, as well as the anaphor. This annotation has four properties: type of anaphoric expression, type of antecedent, a weight manually assigned to the

[1] The topic concept in the studies of Rocha is a noun phrase with the same sense that focus concept in discourse structure theory.

antecedent, and information used for anaphora resolution. The annotated dialogues are used for feeding statistics collectors that generate information for statistical-based anaphora resolution systems. Other manually segmented dialogues are used by Carletta et *al.*(1997) and Byron and Allen (1998) on restricted domains.

As it is shown in Rocha (1998), the use of the dialogue structure information can provide an adequate space for anaphoric accessibility. We also consider that dialogue segments can be detected according to its topic or theme (relevant dialogue entity) scope. And this topic can be proposed according to both frequency and distance of the occurrences. Like Mitkov (1996), the proposed algorithm combines linguistic approaches (restriction and preference systems) with space of anaphoric accessibility information.

Therefore, three spaces of anaphoric accessibility have been defined. Two of them are rigid and depend on the dialogue *interventions*. The third one is variable and is led by the topicality of the dialogue, i.e, what topic or subject the dialogue is about when the anaphor is processed. We present first the general framework of the pronoun resolution algorithm, then describe in detail the proposed approach, with a detailed example. We present then the results of the evaluation we performed on ten Spanish dialogues and discuss the results.

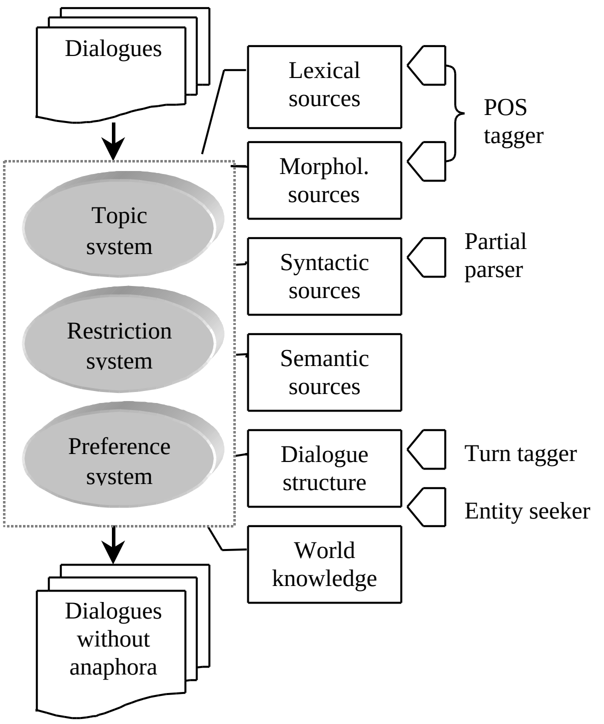

Fig. 1. Global framework for anaphora resolution in dialogues

2. A global framework for anaphora resolution in Spanish dialogues

Figure 1 gives an overview of the framework we propose for pronoun resolution in dialogues. The algorithm uses three systems: a restriction system, a preference system and what we call a *topic* system.

The three systems interacts as follow:

- The **topic system** operates on the defined framework establishing spaces of anaphoric accessibility and provides a list of candidates antecedents. Intuitively, this system marks the boundaries of the parts of the dialogues where the pronoun antecedent must be found.
  The topic system generates dialogue structure information and lists of *topics* extracted from the dialogue. Such topic lists are the potential antecedents and input to the set of restrictions and preferences.
- The **restriction system** contains the rules needed for choosing a possible antecedent for a certain anaphora. Any antecedent that fails in one of the restrictions is rejected.
- The **preference system** is used to disambiguate antecedents when there are several candidates, i.e, when the restriction system has generated more than one antecedent.

Both restrictions and preferences systems are based on morphologic, lexical and syntactic information described in detail in Ferrandez et *al.* (1998).

Using the structure of the dialogue (i.e, the topic system only) as the only information for pronoun resolution, is obviously not enough, as confirmed by the bad results we obtained when excluding the preference and restrictions systems. On the other hand, doing the opposite experiment, i.e, without the dialogue structure information, and applying Ferrández et *al.* (1998) system, that performed well on non-dialogue texts, on dialogues, showed a significant drop of the pronoun resolution performances.

These three systems take input from six information sources:

- **Lexical sources**: They provide information about certain words or groups of words, such as their *implicit verb causality* or *causal valence* as defined by Grober et *al.* (1978). This type of information is used as a preference. For example in *Ann bit Helen because she was better*, the implicit verb causality of the verb *bit*, will prefer *Ann* as the antecedent of *she*. The current system works with word roots and needs the POS tagger output.

- **Morphologic sources**: provide number, gender and person information for the anaphoric expression and the proposed antecedent, using the output of the POS tagger.
- **Syntactic sources**: provide a partial parsing of sentences using the partial parser, Slot Unification Partial Parser (SUPP). Such information allows defining syntactic restrictions between pronouns and candidate antecedents, such as the c-command restriction. It also allows defining preferences such as the syntactic parallelism described in Ferrández et *al.* (1998).
- **Semantic sources**: provides for example semantic information about roles of verbs, pronouns can inherit from the sentence verb and be used then as restrictions on the proposed antecedents. The current implementation of our system does not include such data. However, an addition of such resources is under study in our group.
- **Dialogue structure**: delimits the scope of each dialogue participant and generates a weighed list of *topics* that represent the space of anaphoric accessibility. A detailed description of the algorithm that constructs the topic lists is given is section 3.
- **World knowledge:** It would be useful to represent the world knowledge the listener/reader has on certain situations. This knowledge allows the listener/reader to solve anaphora like in the following example: *The President of the U.S.A. spoke to mass media. Mr. Clinton apologized for making them wait.* In this case, only the world knowledge allows the listener to relate *Mr. Clinton* to *The President of the USA.* Representing the world knowledge is known to be a non-trivial task and it is not represented in our current implementation of the system. Nevertheless, we will leave a place in our global framework to include this information when available.

3. Anaphora resolution algorithm in dialogues

3.1. Definitions

Before presenting the algorithm for anaphora resolution, some definitions are introduced below:

3.1.1. Entity

An entity is defined as a unit of information, which can be an anaphor antecedent. This algorithm only treats pronominal anaphora, and we will treat only anaphors that have a noun phrase as antecedent. Then, syntactically, an entity is a noun phrase.

3.1.2. Participant

A participant is one of the dialogue speakers.

3.1.3. Turn

A turn is a text fragment in which a participant speaks without interruption. According to Gallardo's (1998) definition, a turn can be classified into intervention or continuation turn:

a) **Intervention turn**: where the participant adds information, i.e., occurrence of one or more entities as defined above in 3.1.1. As from now, we will refer to it as **intervention**.
b) **Continuation turn**: where the participant does not add entities. A participant uses this turn in order to encourage the current participant to continue its intervention.

For example, in the next fragment:

```
T1: I have just buy a new car.
T2: Really?
T3: Yeah, I am going to drive it now for
    first time.
```

T1 and T3 turns introduce entities (*a new car* and *it*), so they are interventions. But T2 does not introduce entities, so this is a continuation turn. Therefore, we can consider T1 and T3 as the same intervention interrupted by T2.

3.1.4. Topic

The topic is defined as the most relevant entity in a text fragment. The participant focuses on it and refers to it several times during this fragment.

3.2. Spaces for anaphoric accessibility

As described previously, the topic system provide lists of candidate antecedents to the restriction and preference systems in order to choose the appropriate antecedent. The topic system defines three spaces for anaphoric accessibility the pronoun resolution algorithm uses to look for the correct antecedent:

a) **Current intervention**: is the first space where algorithm looks for an antecedent. It can be easily determined because dialogues are annotated with the turns, using marks like <*Hn*>, where *Hn* indicates the participant who is speaking.
b) **Previous intervention**: if no antecedent has been found in the current intervention, the algorithm looks for an antecedent in the previous intervention.
c) **Relevant topic scope in dialogue**: if no antecedent has been proposed in the spaces above, the algorithm looks for it in the *relevant topic scope* space. This space is determined by the frequency and distance criteria of the entities occurring before the pronoun. It is represented by the list of entities that has been selected by the algorithm using these criteria. For this, the algorithm uses two coefficients:

Cf: *Coefficient of frequency.*
Ci: *Coefficient of in-frequency*

Cf increases the weight of an entity to focus its importance when the entities appear in the current intervention. *Ci* decreases the weight of entities that appeared in the previous interventions, to indicate the loss of importance of those entities. Both parameters are obviously affecting weights of entities to reflect the frequency of the entities in the dialogue and their distance to the current intervention. The entity with the highest weight value will be the most relevant topic of the current intervention, therefore the first favoured candidate antecedent of this whole list. So, the relevant topic scope is the dialogue space where all the relevant entities with a positive weight value (> 0) occur.

We assign the values 10 and 1 units respectively to *Cf* and *Ci*. These values have been proposed experimentally, but a study of possible variations will be needed in order to improve the results, as discussed later on.

Space lists

The spaces defined above are represented using three different lists containing the candidate antecedents:

a) The ***current local entity list (Fa)*** contains the list of entities that appear in the scope of current intervention.
b) The ***previous local entity list (Fa')*** contains the list of entities that appear in the scope of the previous intervention.
c) And, finally, the ***general topic list (F)*** contains the list of entities occurring in the previous interventions of the dialogue and ordered by their weights assigned according to frequency and distribution criteria.

3.3. Algorithm

The three lists introduced above are the basis for the proposed algorithm shown in figure 2.

This algorithm analyses each dialogue in the following way:

1. In each turn, all the sentences are analysed and entities (noun phrases) extracted and stored in *current local entity list (Fa)* (line 5 in Fig.2).
2. For each turn, when the first entity is found and added to *Fa*, the current turn is considered as an *intervention* (see definitions in section 3.1) and the algorithm empties the *previous local entity list (Fa')* when the previous turn is also an *intervention*. In the case of a *continuation turn*, the algorithm keeps *Fa'* entities in order to extend the space of anaphoric accessibility (lines 11 to 18).
3. When an intervention ends, the algorithm incorporates each entity stored in the *current local entity list (Fa)* in the *general topic list (F)*. This operation involves, on one hand, to increase the weight of these entities using *Cf* and, on the other hand, to decrease the weight of entities in *F* and not occurring in *Fa*, using the coefficient of *Ci*. Finally, *Fa* is emptied and saved in the *previous local entity list (Fa')* (lines 20 to 27).
4. During the processing of each sentence, the algorithm checks the occurrence of pronominal anaphors. For each pronoun, it proceeds as follows:
   · The set of restrictions and preferences are applied on the *current local entity list (Fa)* first, then on the *previous local entity list (Fa')* if no antecedent has been proposed from *Fa*.
   · If the pronoun still unresolved, elements of the *general topic list (F)* (lines 6 to 10) are proposed, sorted with the highest weight, and preferences are applied for entities with equal weight. It is important to notice that entities in *F* with a weight value decreased to 0, will not be considered again as candidate antecedents. This shows the strong impact of both *Cf* and *Ci* coefficient values on the scope of the space for anaphoric accessibility.

```
1 Begin
2   Fa = Fa' = F := EMPTY
3   While turns
4     While sentences
5       Fa := Fa+Parse_entities(Sentence)
6       If Search_anaphor
7         Choice_antecedent(Fa)
8         Or Choice_antecedent(Fa')
9         Or Choice_antecedent(F)
10      EndIf
11      If Fa<>EMPTY
12        Kind(turn) :=INTERVENTION;
13        If ( Kind(turn-1)=INTERVENTION or
14           Participant(prev_intervent)
15           <>Participant(turn))
16          Fa' :=EMPTY;
17        EndIf
18      EndIf
19    EndWhile
20    If Fa<>EMPTY /* intervention */
21      F :=incorporate(Fa,F)
22      Fa' := Fa+Fa'
23      Fa :=EMPTY
24      Last_intervention :=turn
25    Else
26      Kind(turn) :=CONTINUATION_TURN;
27    EndIf
28  EndWhile
29 End
```

Fig. 2. Pseudo-code of the algorithm for anaphora resolution.

3.4. An example

In the following we give an example of Spanish dialogue fragment (with its English translation) with a detailed trace of the algorithm steps. *Tnn* marks the turn number to be processed.

In a fruit stall.

T01: <H1> Buenos días.
Good morning.

T02: <H2> Buenos días. ¿Qué deseas?.
Good morning. Can I help you?

T03: <H1> Quiero **manzanas**.
*I want **apples**.*

T04: <H2> ¿De qué **clase** **las** quieres?.
*What kind of **them** do you want?*

T05: <H1> No importa si son buenas.
It does not matter if they are good.

T06: <H2> **Éstas** **las** he recibido **esta mañana**. Son muy buenas.
*I have received **these this morning**. They are very good.*

T07: <H1> Entonces dame de **esas**.
*Then give some me from **those**.*

T08: <H2> ¿Cuántas quieres?.
How many do you want?

T09: <H1> Media docena.
Half a dozen.

T10: <H2> Muy bien, ¿qué más quieres?.
Ok, anything else?

T11: <H1> ¿Tienes **limones**?.
*Do you have **lemons**?*

T12: <H2> Sí, **los** tengo en **una caja** por aquí... Aquí están. ¿**Los** quieres muy verdes?.
*Yes, I have **them** in **a box** near... Here they are. Do you want **them** to be unripe?*

T13: <H1> No... Son para hacer **limonada**. A **mis hijos** **les** encanta.
*No... They are for making **lemonade**. **My sons** enjoy **it** very much.*
T14: <H2> ¿Cuánto quieres?.
How much do you want?
T15: <H1> **Un kilo**.
One kilo.

T16: <H2> **Éstos** te **los** dejo a **buen precio**. Muy bien, ¿algo más?.
*I will give you **these** at **good price**. Very good. Anything else?*

T17: <H1> Nada más, gracias. ¿Cuánto es todo?.
Nothing else, thanks. How much is it?

T18: <H2> 350.

T19: <H1> Ahí tienes. Hasta luego.
Here you are. Bye.

T20: <H2> Hasta luego.
Bye.

The algorithm starts considering turns T01 and T02 as continuation turns due to the lack entities. So, the three lists remain empty until T03. When T03 is processed, (*manzanas*) is considered as a new entity. This entity is included in the *current local entity list (Fa)*:

Fa = [manzanas]

T03 is considered as an intervention. When this intervention finishes all the entities in *Fa* are added to general topic list (*F)* and *Fa* is emptied saving it in *previous local entity list (Fa').* This operation results in the following values for the lists (notice that the pronoun *I* in Spanish is implicit):

Fa = [ ]
Fa' = [manzanas]
F = [(manzanas,10)]

In T04, after including *clase* as new entity, the first anaphor is detected (the pronoun *las*). At this point, the status of the lists is:

Fa = [clase]
Fa' = [manzanas]
F = [(manzanas,10)]

In order to solve the anaphor, the algorithm looks for an antecedent in the *current local entity list (Fa)*. In this case the entity (*clase)* is found, but the restriction system rejects it because of its morphologic and syntactic characteristics. So, the algorithm must look for a possible antecedent in the previous local entity list *Fa'*, in this case (*manzanas*). After applying restrictions, *manzanas* is proposed as an antecedent for the pronoun, and is included as a new entity in *Fa*:

Fa = [clase, manzanas]

After processing the first sentence of T04, *Fa'* is emptied:

Fa' = [ ]

And when T04 ends the lists are modified again:

Fa = [ ]
Fa' = [clase, manzanas]
F = [(manzanas,20),(clase,10)]

T05 is a continuation turn and do not modify the lists. After T06, the weight of *manzanas* increases due to new occurrences of this instance (introduced by pronouns) while the entity *clase* loses it:

Fa = [ ]
Fa' = [manzanas, esta mañana]
F = [(manzanas, 40), (esta mañana,10), (clase, 9)]

The algorithm repeats this process until dialogue ends.

We can remark that anaphora caused by the pronouns *éstos* and *los* in the turn *T16* can be only resolved using entities from the *general topic list F* because local lists *Fa* and *Fa'* do not contain the correct antecedent:

Fa = [ ]
Fa´ = [un kilo]
F = [(manzanas, 45), (limones, 39), (un kilo, 10), (caja, 8), (media docena, 6), (esta mañana, 4), (clase, 3)]

In this case (turn *T16*):

a) *Current local entity list (Fa)* is empty.
b) Restriction system rejects *previous local entity list (Fa').*

c) Finally, restriction and preference systems act over *general topic list (F)* and return *limones* as antecedent.

4. Evaluation of pronoun resolution algorithm

4.1. Results

The algorithm, previously presented, has been implemented using LPA-Prolog and using the input provide by SUPP, see Martínez-Barco et *al.* (1998), and the POS tagger of Xerox. Ten independent Spanish dialogues have been used for the evaluation. In table 1, the different characteristics of the 10 dialogues are described, including the number of sentences, the number of words and the type of dialogue.

# Dialogue	# Sentences	# Words	Type of dialogue
1	163	1376	FDT
2	328	3076	FDT
3	24	545	FDT
4	136	911	FDT
5	62	554	FDT
6	10	100	CDT
7	14	107	CDT
8	13	105	CDT
9	8	63	CDT
10	15	115	CDT
10	773	6952	

Table 1: Dialogue characteristics

We work on two different types of dialogues: transcriptions of free dialogues (FDT) and transcriptions of controlled dialogues (CDT) we created ourselves. Notice that the FDT dialogues[2] are transcribed directly from recorded tapes. These dialogues have a very important amount of incomplete sentences and ellipsis. In this case, the partial parser will process incomplete sentences, and the system will try to extract at least all the completed noun phrases.

In table 2, are given the results obtained by the system. The table contains for each dialogue (D): the total number of pronouns (N), the number of solved pronouns (S) and the number of correctly solved pronouns (C) as well as the values of precision (P), recall (R) and accuracy (A) based on the definitions below:

· *Precision (P)* is the quotient between the number of correctly solved pronouns (C) and the number of solved pronouns (S).

· *Recall (R)* is the quotient between number of correctly solved pronouns (C) and the total number of pronouns (N).

· *Accuracy (A)* is the quotient between S and N

D.	N	S	C	%P	%R	%A
1	11	9	7	78	64	82
2	24	24	16	67	67	100
3	1	1	1	100	100	100
4	13	13	10	77	77	100
5	1	1	1	100	100	100
6	9	9	7	78	78	100
7	4	4	3	75	75	100
8	5	5	5	100	100	100
9	5	5	5	100	100	100
10	7	7	5	71	71	100
	80	78	60	77	75	98

Table 2: Output system results

In table 3, we present different antecedent lists (Fa, Fa', F). Columns 2-4 show the number of real antecedents that need to be found by the system in each of the lists Fa, Fa' and F, then the columns 5-7 the number of those antecedents that have been correctly identified by our system.

	Antecedents			Correctly Antec. Chosen		
D.	Fa	Fa'	F	Fa	Fa'	F
1	5	2	2	4	2	0
2	13	7	4	11	5	0
3	0	1	0	0	1	0
4	5	4	4	5	4	1
5	1	0	0	1	0	0
6	3	4	2	3	3	1
7	0	4	0	0	3	0
8	0	5	0	0	5	0
9	1	4	0	1	4	0
10	0	6	1	0	5	0
	28	37	13	25	32	2

Table 3: Antecedents

Finally, table 4 shows a classification of the errors produced by the system. We distinguish 6 different reasons for the failure of the system, although there might depend on each others:

	ERRORS					
D.	T	CE	PR	SR	W	O
1	0	1	0	0	2	0
2	2	3	0	1	2	1
3	0	0	0	0	0	0
4	0	0	2	0	1	0
5	0	0	0	0	0	0
6	0	0	1	1	0	0
7	0	0	1	0	0	0
8	0	0	0	0	0	0
9	0	0	0	0	0	0
10	0	0	1	0	0	1
	2	4	5	2	5	2

Table 4: Error system

· The *Tagger* (T): error due to an incomplete tag set. In fact, the two fails detected are due to the fact our algorithm chooses as an antecedent for

[2] Free Dialogues Transcriptions have been provided by the Computational Linguistic Laboratory of the Autonomous University of Madrid

el (singular masculine pronoun, *he*) the name of a city, *Almería*, that has a feminine gender. Our tag set does not provide the gender of place-names.

- *Carried error* (CE): it is when a pronoun is incorrectly solved causing a bad side-effect on the weight values of entity, for further pronoun resolutions.
- *Poor restriction* (PR): when an incorrect antecedent should be eliminated from the list but is not by the current restrictions.
- *Strong restriction* (SR): when restrictions eliminate a correct antecedent from the list.
- *Weight* (W): when the correct antecedent does not have the higher weight in the list.
- *Others preferences* (O), when wrong preferences are applied.

4.2. Discussion

We have obtained a 77% of precision, which is a good score considering the fact that the algorithm uses partial parsing information and that it does not use semantic information.

Nevertheless, as we can see in table 3, our system captures most of the current and previous intervention spaces (25 over 28 and 32 over 37 anaphors were correctly solved respectively in these spaces), while it only captures 2 out of 13 of the relevant topic scope. This suggest a stronger definition of the relevant topic scope.

On the other hand, table 4 shows the different causes of failures detected in our system. Most frequent ones are the use of poor restrictions, bad weights assigned and carried errors. We have analyzed these problems and we can propose the following improvements:

- The use of semantic sources is needed in order to improve the restriction system with stronger filters.
- The problem of bad weights assignment suggest to study in detail the possible variations of the values to assign to the c*oefficient of frequency (Cf)* and the *coefficient of in-frequency (Ci).* These values have been proposed experimentally and further evaluations with different values are needed.

We have estimated that without the tagger errors and without carried errors, we would obtain around 85% and 83% of precision and recall respectively.

Conclusion

We have presented an algorithm for anaphora resolution in Spanish dialogues mixing linguistic information with dialogue structure information. The presented work represents one of the first works done in the study of pronominal anaphora resolution in Spanish dialogues, we are not aware of any other similar word for dialogues in the Spanish language.

We have performed a detailed evaluation of this algorithm using ten Spanish dialogues. These dialogues have been selected from different sources and they have different characteristics. The obtained results, 77% of precision and 75% of recall, respectively, are really encouraging. We discussed in details the obtained results, the main reason of failure and how to improve the algorithm. Finally, we performed another evaluation showing that the results can be significantly improved up to 85% of precision and 83% of recall without the tagger and carried errors.

Acknowledgements

This paper has been supported by the Comisión Interministerial de Ciencia y Tecnología with project number TIC97 - 0671-C02 - 02.

References

Azzam S., Humphrey K. and Gaizauskas R. (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,* Montreal, Canada. pp 74-78

Byron D.K and Allen J.F. (1998) Resolving Demonstrative Anaphora in the TRAINS93 Corpus. In *Proceedings of Second Colloquium on Discourse Anaphora and Anaphor Resolution (DAARC2)*, **11**. Simon Botley & Tony McEnery Eds. (Lancaster, 1998). pp 68-81.

Byron, D. and Stent, A. (1998) A Preliminary Model of Centering in Dialog. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, COLING-ACL*. Montreal. Canada.

Carletta J., Isard A., Isard S., Kowtko J.C., Doherty-Sneddon G. and Anderson A.H. (1997) The Reliability of a Dialog Structure Coding Scheme. *Computational Linguistics,* **23**, nº 1. March, 1997. pp 13-32.

Ferrández A., Palomar M. and Moreno L. (1998) Anaphor resolution in unrestricted texts with partial parsing. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics, COLING-ACL,* (Montreal, Canada). pp 385-391.

Gallardo B. (1998) *Comentario de textos conversacionales*. Colección Comentario de Textos. Arco Libros, S.L., Madrid.

Grober, Beardsley and Camarazza. (1978). *The implicit verb causality or causal valence.*

Grosz B and Sidner C. (1986) Attentions, intentions and the structure of discourse. *Computational* Linguistics, **12**, pp 175-204.

Grosz B., Joshi A. and Weinstein, S. (1995) Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics*, **21**. pp 203-225.

Manabu O. and Kouji T. (1996) Zero Pronoun Resolution in Japanese Discourse Based on Centering Theory. In *Proceedings of the 16th International Conference on Computational Linguistics. COLING'96*, **II**, (Copenhagen, Denmark), pp 871-876.

Martínez-Barco P., Peral J., Ferrández A., Moreno L. and Palomar M. (1998) Analizador Parcial SUPP. In *Proceedings of the 6th Iberoamerican Conference on Artificial Intelligence, IBERAMIA'98* (Lisbon, Portugal). pp 329-341.

Mitkov R. (1996) Anaphor resolution: a combination of linguistic and statistical approaches. *In Proceedings of the Discourse Anaphora And Resolution Colloquium, DAARC96.*

Rocha M.A.E. (1998) *A corpus-based study of* anaphora *in dialogs in English and Portuguese.* PhD Thesis. University of Sussex.

Walker M.A. (1997) Centering, Anaphora Resolution, and Discourse Structure. *Centering In Discourse*, eds. Walker, M.A.; Joshi, A.K.; Prince, E.F., Oxford University Press.

Webber, B. (1988) Discourse deixis: Reference to discourse segments. In *Proceedings of the 26th annual meeting of the Association for Computational Linguistics (ACL'88).* pp. 113-122. 1988.