



HAL
open science

How useful are similarity word lists for indirect anaphora resolution?

Susanne Salmon-Alt, Renata Vieira, Caroline Gasperin

► **To cite this version:**

Susanne Salmon-Alt, Renata Vieira, Caroline Gasperin. How useful are similarity word lists for indirect anaphora resolution?. Actes : 5th International Conference on Discourse Anaphora and Anaphor Resolution, Furnas, Portugal (September 23 - 24, 2004)., 2005. halshs-00005020

HAL Id: halshs-00005020

<https://shs.hal.science/halshs-00005020>

Submitted on 17 Oct 2005

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

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

How useful are similarity word lists for indirect anaphora resolution?

Caroline Gasperin* Susanne Salmon-Alt** Renata Vieira*

*PIPCA - Unisinos
São Leopoldo, Brazil
{caroline, renata}@exatas.unisinos.br
**ATILF - CNRS
Nancy, France
Susanne.Salmon-Alt@atilf.fr

Abstract

This work tests the use of word similarity lists for anaphora resolution in Portuguese and French corpora. We applied an automatic lexical acquisition technique over parsed texts to identify semantically similar words. We then made use of this lexical knowledge to resolve coreferential definite descriptions where the head noun of the anaphor is different from the head noun of its antecedent, which we call indirect anaphora. We compare the results in both languages with a baseline algorithm. Our proposed technique achieved an F-measure of 0.26 for Portuguese and 0.09 for French.

1. Introduction

Coreference has been defined by (van Deemter & Kibble, 2000) as the relation holding between linguistic expressions that refer to the same extralinguistic entity. Anaphora is a slightly different discourse relation. In an anaphoric relation the interpretation of an expression is dependent on previous expressions within the same discourse (in various ways). Therefore, an anaphoric relation may be coreferential or not. An expression may or may not be anaphoric in the strict sense that its interpretation is only possible on the basis of the antecedent, as it is in general the case of pronouns in written discourse. On the other hand, it might be only coreferential, in the sense that the entity has been mentioned before in the text.

We focus on expressions that are anaphoric and coreferential, and we furthermore limit our research to indirect cases of these phenomena, that is, when the antecedent's and the anaphor's head noun are not the same but are semantically related. This classification, also adopted in our previous work (Vieira et al. 2002, Salmon-Alt & Vieira, 2002) was based on the analyses of English texts presented in (Poesio & Vieira, 1998), with the difference that we divided the bridging class of their analyses into two different classes, separating coreferential (*indirect anaphora*) and non-coreferential (*other anaphora*) cases. The reason for concentrating on indirect anaphora is that they pose an interesting computational problem, since they are heavily dependent on common or lexical knowledge to be interpreted. From all bridging cases, they are those which could be clearly classified as anaphora (differently from other non coreferential anaphora) in the corpus annotation. Tables 1 and 2 show some examples of indirect anaphora from the Portuguese and French corpora.

Table 1: Examples of indirect anaphora in Portuguese

	Antecedent	Anaphor
1	a denúncia de que ... (<i>the denunciation that ...</i>)	a acusação (<i>the accusation</i>)
2	São_Paulo	a cidade (<i>the city</i>)
3	A idéia de que ... (<i>the idea that ...</i>)	a proposta (<i>the proposal</i>)

Table 2: Examples of indirect anaphora in French

	Antecedent	Anaphor
1	une armée divisée (<i>a divided army</i>)	les soldats (<i>the soldiers</i>)
2	Barcelone	la ville (<i>the city</i>)
3	des crimes ... (<i>the crimes</i>)	les faits (<i>the facts</i>)

This work investigates the use of word similarity lists for treating indirect anaphora. We applied a lexical acquisition technique (Gasperin, 2001; Gasperin et al. 2001) on Portuguese and French parsed corpora to automatically identify semantically similar words. The acquired lexical knowledge is then consulted for the resolution process.

This paper is organized as follows: the next section details the tools and techniques used for the construction of our lexical resource. Section 3 presents our heuristics for solving the indirect anaphors on the basis of this resource and describes the experiments we did to evaluate the heuristics. Section 4 discusses our results with respect to related work. In section 5, we detail some interesting points concerned with particular parameters of our experiments. Finally, section 6 presents our concluding comments.

2. Acquisition of similar word lists

Our lexical resource consists of lists of semantically related common nouns. These lists are constructed automatically by a syntax-based knowledge-poor technique. The technique used is described in (Gasperin, 2001; Gasperin et al. 2001), and it is an extension of the technique presented in (Grefenstette, 1994).

To resume, this technique consists in extracting specific syntactic contexts for every noun in the parsed corpus and then applying a similarity measure, the weighted Jaccard measure, to compare the common nouns by the contexts they have in common (the more contexts they share, the more they are similar). By syntactic context we mean any word that establishes a syntactic relation with a given noun in the corpus. An example of one kind of syntactic context considered is subject/verb, meaning that two nouns that occur as subject of the same verb share this context. We used 6 different sorts of

syntactic contexts: verb/direct-object, verb/indirect-object, subject/verb, subject/object, adjective/noun, and noun/preposition-noun. A global and local weight is assigned to each context: the first is related to the context frequency in the corpus, and the second is related to its frequency as a context of the noun in focus.

To generate the similarity lists for Portuguese we utilized a 1,400,000-words corpus from the Brazilian newspaper *Folha de São Paulo*, containing news about different subjects (sports, economics, computers, culture, etc.). The corpus was parsed by the Portuguese parser PALAVRAS (Bick, 2000). For French, we utilized a 1,000,000-words corpus from the French newspaper 'Le Monde' to generate the similarity lists. This corpus was firstly POS-tagged by the Decision Tree Tagger (Schmid, 1994) and then parsed by the French parser FrAG (Bick, 2003). Whereas the Portuguese corpus includes the set of texts that was hand-annotated with coreference information in previous work, it was not the case for French (Vieira et al. 2002, Salmon-Alt & Vieira, 2002).

As output, we have a list of the most similar nouns to each common noun in the corpus, ordered by the similarity value. The generated lists can be quite long; we then decided to keep just the 15 words that are most similar to the word in focus, according to the calculated similarity values. This decision leads us to the loss of symmetry in a similarity relation. We present examples of similarity lists in Figure 1.

patron (boss)	acusação (accusation)
actionnaire (shareholder)	denúncia (denunciation)
holding (holding)	escândalo (scandal)
filiale (subsidiary company)	crime (crime)
propriétaire (owner)	pedido (demand)
maire (major)	declaração (declaration)
représentant (representative)	proposta (proposal)
patronat (employers)	notícia (news)
constructeur (constructor)	carta (letter)
magnat (magnate)	lista (list)
roman (novel)	cargo (post)
partenaire (partner)	ataque (attack)
journal (POS error)	arma (gun)
quotidien (newspaper)	caso (case)
associé (associate)	impressão (impression)
assureur (insurer)	reclamação (complain)

Figure 1: Similarity list for French and Portuguese

The similarity lists can contain any kind of lexical relation between the entry noun and associated words:

- quasi synonymy : *accusation* – *denunciation*
- hyponymy : *accusation* – *declaration*
- (loose) meronymy : *boss* – *subsidiary company*
- subset relations : *boss* – *employers*

But more generally, there seems to be quite loose semantic relations, grouping together words of the same semantic field and/or the same semantic type: For French, for instance, there are 10 words related to the semantic field of entrepreneurship (*holding*, *subsidiary company*, *owner* etc.) and nine words which share with *boss* the property of being animated and denoting the roles (*owner*, *partner*, *associate* etc.). However, some of them are very loosely semantically related words (*partner*), corpus specific collocations (*newspaper*).

Noise in the lists is either due to part of speech or parsing errors (*journal*, for instance, is a bad lemmatization of *journaux*) or to the fact that the similarity lists for less frequent words were based on few syntactic contexts they shared along the corpus. One of the open issues is the qualitative evaluation of these similarity lists. Since the expected results of the acquisition technique we used is context dependent *word similarity* in a broad sense, it is difficult to evaluate them extensively against other resources (when available) or human judgment.

However, we believe that this technique has two main advantages in the context of anaphora resolution. The first is the possibility of having an automatically built lexical resource. This is an important issue for languages that lack good quality and/or wide coverage lexical resources for NLP such as *WordNet* for English. This is in fact the case for many languages except English, in particular for French and Portuguese.

The second advantage is the possibility of having a corpus-tuned lexical resource. This resource closely reflects the semantic relations present in the corpus used to create the lists. Thus, the use of similarity lists for resolving anaphora might be thought of as being more suitable than a perhaps (insufficient) generic lexical database (e.g. French *Wordnet*), since it focuses on the semantic relations between the terms that appear in the corpus without considering extra meanings that some words could have.

In order to test the relevance of this second hypothesis, we decided to run the lexical acquisition in a slightly different way for each of the languages studied: the French acquisition is an exogenous approach, in the sense that the test corpus used for anaphora resolution has not been included in the acquisition corpus for similarity lists. We did however use the same text genre (newspaper *Le Monde*). For Portuguese, we did use an endogenous strategy: not only was the text genre (newspaper *Folha de São Paulo*) the same, but the test corpus for anaphora resolution was also included in the acquisition corpus.

3. Similar word lists and indirect anaphora

3.1. Test corpus

The French coreference corpus consists of a selection of newspaper articles of *Le Monde* - different from those used for the acquisition of similarity lists - about various topics. It contains roughly 9000 words and 741 definite descriptions annotated with coreference links by two annotators. Both considered about 30% of the definite descriptions as coreferential, but there was an important difference with respect to the direct vs. indirect classification. Therefore, we chose for this study only those definite descriptions which had been annotated by both annotators as an indirect anaphor. This was the case for 78 descriptions. Forty-nine of them, having at least one antecedent which was not a named entity, were chosen for our study. For these 49 anaphora, the annotators highly agreed on the antecedents: in only a few cases (12%) was there a difference between the two annotations. In this case, we considered both possibilities as key for the output of our system.

For Portuguese, the corpus contains around 6000 words from 24 news texts of six different newspaper

sections of *Folha de São Paulo*. From a set of 680 definite descriptions, we filtered out 95 indirect anaphora. From these, we considered 57 cases not related to a proper noun as an antecedent for this study. As for French, the manual annotation was done by two annotators. The classification agreement for the classes coreferent and non-coreferent was $K=0.76$ (Kappa measure), and the agreement for classifying coreferent cases as direct and indirect was $K=0.57$. This was a much higher agreement than for the distinction between other anaphora and discourse new from the non-coreferent cases ($K=0.29$).

3.2. Usefulness of similar word lists for resolving indirect anaphora

3.2.1. Possibilities of using similarity lists for indirect anaphora resolution

The similarity lists can be used in different ways for the resolution of indirect anaphora. If one consider:

- H_{ana} as the head noun of the anaphor;
- H_{can_i} as the head noun of the antecedent candidate i ;
- L_{ana} as the anaphor's list of similar nouns;
- L_{can_i} as the list of similar nouns for the candidate i ;

then, H_{can_i} is considered as the antecedent of H_{ana} if

- (1) $H_{can_i} \in L_{ana}$ or
- (2) $H_{ana} \in L_{can_i}$ or
- (3) $H_{ana} \ni L_x \in H_{can_i}$
- (4) $L_{ana} \ni H_x \in L_{can_i}$

Heuristic (1) is the most intuitive, expecting the antecedent among the similar nouns in the entry for the anaphor. Since the limitation of the number of similar words to 15 causes the loss of symmetry of the similarity relation, strategy (2) takes the opposite way, looking for the anaphor in the similar noun list of the candidate. In (3), we enlarge the search strategy and check if both anaphor and antecedent head nouns are present in a same similarity list, and in (4), we check whether there is a third word which is similar to both the anaphor and the antecedent.

With regards to the expected semantic relatedness between H_{ana} and H_{can_i} , we consider (1) > (2) > (3) > (4). However, this order is rather arbitrary, following only what we believe is the more natural way to search for the similarity relations between two words. We assume that the similarity relations from (1) and (2) are stronger than in (3) and (4), since the nouns had to share many syntactic contexts to be able to be among the first 15 similar nouns of a noun, according to the lexical acquisition technique presented in Section 2. Strategies 3 and 4 were included here as a possibility to recover the similarity cases lost by the pruning of the lists.

3.2.2. Upper performance limit of similar lists

In a first experiment, we evaluated in how many cases both the head noun of the anaphor and the head noun of the correct antecedent could be related by applying one of the previous search heuristics. This score indicates the

upper performance limit for resolving indirect anaphora in our test corpus on the base of similarity lists. Table 3 shows the results for French and Portuguese corpora.

Table 3: Upper performance limit

	#	%
French	12	24,5
Portuguese	33	57,8

The results show that similarity lists indeed contain some of the relevant information for indirect anaphora resolution. The difference of 20 points between French and Portuguese is likely to be due to the difference in the acquisition of the similarity data for the two languages (exogenous vs. endogenous) and is a strong indicator for using the latter approach whenever possible. For French (exogenous acquisition), we observe as an interesting point that the solved cases all involve a strict and context independent synonymy between the concerned head nouns (*peine-condamnation*, *arrêt-panne*, *soldat-militaire*, *absence-déficit*). We compared this result to lexical knowledge about synonymy relations encoded in a good French language dictionary (*Le Grand Robert*) and found, surprisingly, that only one more case could be resolved on the basis of traditional dictionary look up.

As a conclusion, this would mean that even exogenous similarity lists perform almost as well as traditional human built lexical resources, at least for the evaluation of maximum recall, without taking into account the issue of precision (see section 3.2.3). This is probably due to the fact that most of the indirect anaphora in the test corpus rely on contextual rather than on strict lexical relations. As one would expect, those relations are not likely to be encoded in traditional dictionaries, but more surprisingly, they are also missing in the similarity lists.

Interestingly, when considering an endogenous situation (as for Portuguese, where our test corpus is part of the corpus used to generate the similarity lists), very context-dependent similarity cases could be found. For instance, we have the noun *micro* (in micro computer) as similar to *um 386 DX , 40 MHz , 8 Mbytes de RAM*. It is very unlikely that this relation could be found in a dictionary. Also, an endogenous approach could be valid when treating proper names, since dictionaries do not contain named entities.

We next evaluated the efficiency of our search strategies. We counted how many cases each of the heuristics was able to solve. Table 4 shows these results for French and Portuguese. For both languages, strategy 3 was the most useful, being able to resolve almost all the cases. Since we observed that most of the lists that contain both anaphor and antecedent head nouns are lists of low frequent nouns in the corpus, we interpret the fact that strategy 3 can treat more cases since it is the one that deals best with the nouns with average frequency in the corpus. This strategy does not require the nouns to be frequent enough to be in the top positions of the list of the high and medium frequent nouns. At the same time both anaphor and antecedent head nouns can be part of the list of a low frequency third noun which is useless by itself, since it is rare and probably not used in anaphora cases).

However, as shown in Table 5, the combination of all strategies is useful, at least for Portuguese. For French, the

numbers are too low for drawing conclusions about the usefulness of strategies without any application.

Table 4: Comparing Strategies

Strategy	French	%	Portuguese	%
1	5	41,7	11	33,3
2	4	33,3	17	51,1
3	11	91,7	31	93,9
4	7	58,3	23	69,6
Cumulated	12	100	33	100

Table 5: Combining Strategies

Strategy	1	2	3	4	Total
French	5	+0	+7	+0	12
Portuguese	11	+6	+15	+1	33

3.2.3. Automatic resolution

In order to evaluate the usefulness of similarity word lists in the context of automatic resolution of indirect anaphora (evaluating recall, precision and F-measure), we implemented the same resolution heuristics for French and Portuguese and run them over our test corpora. All input (definite NPs with a common head noun classified as indirect anaphor) and output data (antecedents for these NPs) were represented in XML format, and the resolution and evaluation strategies were implemented by XSL style sheets. As possible candidates for antecedents, we considered all full NPs with a common head noun, preceding the anaphor in the same text. The principal parameters for the algorithm are given in Figure 2.

```

For each NP that is an indirect anaphor
look for a:
- preceding NP
- starting with the most recent
  candidate NP
- considering the current sentence
- stopping at a 5-sentence window
Fulfilling our search strategies in the
following order
- (1)> (2)> (3)> (4).
Link the first selected candidate with
the anaphor. If there is no suitable
antecedent, leave the anaphor
unresolved.

```

Figure 2: Algorithm for solving indirect anaphora

3.2.4. Evaluation of the results

We implemented a baseline algorithm to be able to compare our results against it. The baseline algorithm selects simply the most recent NP (without any agreement constraints) as antecedent for the anaphor. Since an antecedent is always found, the recall, precision and F-measure have the same value. In Table 6, we present the recall, precision and F-measure values for our strategy comparing them to the values for the baseline algorithm.

For French as for Portuguese, the baseline scores are based on a very few right responses (5 and 4, respectively). However, the baseline algorithm curiously worked three times better for French than for Portuguese.

Comparing the output of our algorithm using similarity lists to the baseline, the most interesting result is that the scores are below the baseline for French, but much better for Portuguese. Indeed, the score for the F-measure is almost three times better for Portuguese than for French. As for the results reported in the previous section, we believe that this difference must be explained mainly by the fact that we used different acquisition strategies (exogenous vs. endogenous) for these two languages.

Table 6: Baseline comparison

		R	P	F
French	Baseline	11%		
	Simlist	8%	10%	9%
Port	Baseline	3%		
	Simlist	23%	31%	26%

We also investigated the utility of each of the four strategies. The following tables compare the application of each strategy for French and Portuguese.

For both languages, search strategy 3 was the one that caused most of the false positive cases. This is the same strategy that was able to solve most cases according to the upper performance study presented in Section 3.2.2. In fact, in 80% of the false positive cases for strategy 3, the correct antecedent was found but then lost because a closer one was finally selected. The phenomenon of having two semantically related head nouns in a same similarity list is actually more frequent than expected and brings noise to the resolution process.

Table 7: French results per strategy

Strategy	Right positive	False positive	Total
1	0	2	2
2	1	4	5
3	0	22	22
4	3	8	11
Total	4	36	40

Table 8: Portuguese results per strategy

Strategy	Right positive	False positive	Total
1	4	5	9
2	4	1	5
3	4	15	19
4	1	8	9
Total	13	29	42

3.2.5. Error analysis

As shown in section 3.2.2., the most important problem was silence, i.e. pairs of anaphor-antecedent head nouns that were not in the similar lists. This problem was then combined with false positive cases. Most of the false positive cases were due to one of the two following causes:

(1) The right antecedent was not in the lists, therefore it could not be found, but other wrong antecedents were retrieved. For example, in

... meu **amigo** Ives Gandra da Silva Martins escreveu para esse jornal ... o **conselheiro** Ives (my friend Ives Gandra da Silva Martins wrote to this newspaper ... the councilor Ives)

two more candidates head-nouns are similar words to the head noun of the anaphor *conselheiro* (*councilor*): *arquitecto* (*architect*) and *consultor* (*consultant*), but not *amigo* (*friend*).

(2) The right antecedent was in the lists, but another wrong antecedent was taken before, because of proximity to the anaphor, as in the French example

Cette année, le **déficit** est de 20% à 90% dans plus des deux tiers du pays ... cela entraîne des *conséquences* dramatiques pour les récoltes de mousson ... L' **absence** de pluie ... (This year, the deficit is 20 to 90% for more than two thirds of the country ... this brings dramatical consequences for the harvest ... the lack of rain ...)

Here, the correct antecedent to *l'absence* (*lack*) is *déficit* (*deficit*). It is present in *absence*'s similarity list, but this is also the case for the closer candidate *conséquence* (*consequence*), which has been chosen as antecedent.

4. Related work

An evaluation of the use of WordNet for treating bridging descriptions is presented in (Poesio, Vieira & Teufel, 1997). This evaluation considers 204 bridging descriptions: for 12, the antecedent and the anaphor are in a synonymy relation, for 14 of them in a hypernymy relation. Those relations were expected to be found in

WordNet. This was the case for only 33,3% of synonymy and 57,1% hypernymy cases.

The technique presented in (Schulte im Walde, 1997) based on lexical acquisition from the British National Corpus was evaluated against the same cases in (Poesio, Vieira & Teufel, 1997). For synonymy, hypernymy and meronymy, it was reported that 22% of the 38 cases were resolved.

With a slightly different scope (bridging anaphora concerned with meronymy), Poesio et al. (2002) improved the resolution of meronymy by the inclusion of syntactic patterns for a lexical knowledge acquisition strategy. This results in 66% of the meronymy cases being resolved. Finally, Bunescu (2003) reports a precision rate of 53% and a recall rate of 22.7%. for a method that resolves non-coreferential associative anaphora.

Table 9 shows a comparison between the work done here and the related work referred above. These results are presented to give an idea of the results reported in different works, and are not a comparison of the techniques themselves, since the corpora and the anaphora cases treated differ in each experiment.

Table 9: Related work

	Poesio et al., 1997	Schulte im Walde, 1997	Poesio et al., 2002	Bunescu, 2003	Gasperin et al, 2004	
Semantic relation	Synonymy, Hypernymy	synonymy, hypernymy, meronymy	meronymy	associative (non-coreferential)	any relation involved in indirect coreferential noun anaphora	
Language	English	English	English	English	French	Portuguese
Resource	WordNet	acquired	WordNet and acquired	acquired	acquired (exogenous)	acquired (endogenous)
Precision	-	25-33% , 15%, 18-36%	72%	53%	10%	31%
Recall	33% , 57%	25-33%, 15%, 18-36%	66%	22.7%	8%	23%

5. Discussion

This section discusses results of some additional investigations on particular issues, conducted on either French or Portuguese data.

The first point pertains to the role of gender or number agreement for indirect anaphora resolution. Since our basic algorithm does not rely on any agreement rules, we investigated whether such constraints would be able to enhance the performance of our system. Whereas gender agreement was not relevant, number agreement turned out to be a very useful constraint. For French, we obtained an F-measure of 0.19, which is almost twice the score of the heuristics without number agreement. These results are linguistically coherent, since indirect anaphora refer to the same referent (which is generally not expected to be subject of cardinality variation along the discourse), but via different head nouns whose genders are generally linguistically fixed.

Another question is the comparison of our results based on automatically acquired similarity lists with other types of lexical resources. For the French data, we therefore did the same experiments, but replacing the similarity lists by (a) the French part of *EuroWordNet* and

(b) an electronic version of synonyms extracted from a high quality French print dictionary (*Le Grand Robert*). As an interesting result, *EuroWordNet* leads to an F-measure of 0.08, which means worse results than our similarity lists (which are themselves worse than the baseline). This is mainly due to the insufficient coverage of this resource (resulting in a very low recall rates). Less surprisingly, the results obtained with the dictionary synonyms are much better: 0.21 for recall and 0.39 for precision, which leads to an F-measure of 0.27. It is interesting to note here that this score is very close to the one obtained for Portuguese data with endogenous similarity lists (0.26). These results suggest that under certain conditions (availability of big parsed corpora and endogenous acquisition strategy), similarity list do not perform worse than traditional lexical resources.

Furthermore, for Portuguese, we analyzed our results with respect to the lexical relation holding between the anaphor and its antecedent. Out of the 57 cases annotated as indirect anaphora, the relation was synonymy for 14 of them and hyponymy for the remaining 43. For the synonymy cases 64% could be resolved by the proposed heuristics and for the hyponymy cases the number was 28%. For these cases 14% and 35% resulted in false

positives respectively. This confirms our initial intuition about the greater utility of similarity lists for synonymy than for other relations.

Finally, for Portuguese, we also did an experiment including the proper names of the corpus. From the 95 definite descriptions classified as indirect in our evaluation corpus, 38 cases had a proper name as anaphor or antecedent. We regenerated the similarity lists considering the proper names, and each list is then composed by other nouns and proper names. In this experiment, we could solve more cases, but the precision value decreased. We observed that the generated lists including proper names are noisier, mainly because proper names are in general less frequent than common nouns - the lists became more heterogeneous, including more non-semantically related words. When looking at the cases that contained a proper name, for 50% of them we got false positive antecedents and 50% remained unresolved. This means that none of the cases including proper names could be resolved, but does not mean they had no influence on other nouns' similarity lists.

6. Concluding remarks

We tested the usefulness of automatically acquired word similarity lists for resolving indirect anaphora with a common noun head in Portuguese and French newspaper texts. Combining a basic anaphora resolution algorithm based on recency and a 5-sentence search window with four different search strategies on similarity lists, we got an F-measure of 0.09 for French and 0.26 for Portuguese. The difference between these scores is mainly due to a variant in the acquisition procedure: the French lists were acquired on a corpus excluding the test data, whereas the Portuguese lists did not. This argues strongly in favor of an endogenous acquisition strategy whenever possible.

It is difficult to compare the results to related work, especially because of the wide range of what has been understood as bridging anaphora. However, the recall we obtained with an endogenous approach (Portuguese) is globally comparable to the recall reported in previous work using automatic acquisition techniques for solving associative anaphora. More surprisingly, our results, even for an exogenous approach (French), are not worse than results obtained with *EuroWordNet*.

As future work, we intend to take advantage of the similarity values (rather than cutting the lists arbitrarily at 15 words) and to vary the combination of the four search heuristics for the similarity lists. We also plan to use a weighting scheme to balance the distance of an anaphor from its candidate and the search strategy. Other parts of the NP besides head noun could also be taken into account. A more extensive investigation on the basis of larger corpora for the lists generation and with larger numbers of indirect anaphora examples is required in order to verify the initial but promising results we found in this study.

Acknowledgements

We would like to thank CNPq (Brazil) and INRIA and CNRS (France) for their financial support. Thanks also to our annotators, to Eckhard Bick who parsed our corpora and to Paula Chesley for checking the final version.

7. References

- Bean, D. and E. Riloff, 1999. Corpus-based identification of non-anaphoric noun phrases. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL-99)*.
- Bick, E., 2000. *The Parsing System PALAVRAS: Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. Ph.D. thesis, Århus University, Århus.
- Bick, E., 2003. A CG & PSG Hybrid Approach to Automatic Corpus Annotation. In *Proceedings of SproLaC2003*, pp.1-12, Lancaster.
- Bunescu, R., 2003. Associative anaphora resolution: A web-based approach. In *Proceedings of EACL 2003 - Workshop on The Computational Treatment of Anaphora*, Budapest.
- Gasperin, C., P. Gamallo, A. Agustini, G. Lopes and V. Lima, 2001. Using syntactic contexts for measuring word similarity. In *Proceedings of the Workshop on Semantic Knowledge Acquisition and Categorisation*, Helsinki, Finland.
- Gasperin, C., 2001. Extração automática de relações semânticas a partir de relações sintáticas. Master's thesis, PUCRS, Porto Alegre.
- Grefenstette, G., 1994. *Explorations in Automatic Thesaurus Discovery*. Kluwer Academic Publishers, USA.
- Poesio, M. and R. Vieira, 1998. A corpus-based investigation of definite description use. *Computational Linguistics*, 24(2):183-216.
- Poesio, M., R. Vieira and S. Teufel, 1997. Resolving bridging descriptions in unrestricted texts. In *Proceedings of the Practical, Robust, Anaphora Resolution for Unrestricted Texts, Workshop on Operational Factors*, Madrid.
- Poesio, M., T. Ishikawa, S. Schulte im Walde and R. Vieira, 2002. Acquiring lexical knowledge for anaphora resolution. In *Proceedings of LREC 2002*, Las Palmas De Gran Canaria.
- Salmon-Alt, S. and R. Vieira, 2002. Nominal expressions in multilingual corpora: Definites and demonstratives. In *Proceedings of the LREC 2002*, Las Palmas de Gran Canaria.
- Schmid, H., 1994. Probabilistic Part-of-Speech Tagging Using Decision Trees. In *Proceedings of the International Conference on New Methods in Language Processing*, Manchester.
- Schulte im Walde, S., 1997. *Resolving Bridging Descriptions in High-Dimensional Space*. Master's thesis, Institut für Maschinelle Sprachverarbeitung, University of Stuttgart, and Center for Cognitive Science, University of Edinburgh.
- van Deemter, K. and R. Kibble, 2000. On coreferring: Coreference in muc and related annotation schemes. *Computational Linguistics*, 26(4).
- Vieira, R., S. Salmon-Alt and E. Schang, 2002. Multilingual corpora annotation for processing definite descriptions. In *Proceedings of the PorTAL 2002*, Faro.
- Vieira, R., C. Gasperin and R. Goulart, 2003. From manual to automatic annotation of coreference. In *Proceedings of the International Symposium on Reference Resolution and Its Applications to Question Answering and Summarization*, Venice.