Automatic Coreference Annotation in Basque

Iakes Goenaga, Olatz Arregi, Klara Ceberio,
Arantza Díaz de Ilarraza and Amane Jimeno

University of the Basque Country UPV/EHU

iakesg@gmail.com

Abstract

This paper presents a hybrid system for annotating nominal and pronominal coreferences by combining ML and rule-based methods. The system automatically annotates different types of coreferences; the results are then verified and corrected manually by linguists. The system provides automatically generated suggestions and a framework for easing the manual portion of the annotation process. This facilitates the creation of a broader annotated corpus, which can then be used to reiteratively improve our ML and rule-based techniques.

1 Introduction

Coreference resolution task is crucial in natural language processing applications like Information Extraction, Question Answering or Machine Translation. Machine learning techniques as well as rule-based systems have been shown to perform well at resolving this task. Though machine-learning methods tend to dominate, in the CoNLL-2011 Shared Task[1], the best results were obtained by a rule-based system (Stanford's Multi-Pass Sieve Coreference Resolution System [13]).

Supervised machine learning requires a large amount of training data, and the spread of machine learning approaches has been significantly aided by the public availability of annotated corpora produced by the 6th and 7th Message Understanding Conferences (MUC-6, 1995 and MUC-7, 1998) [17, 18], the ACE program [9], and the GNOME project [22]. In the case of minority and lesser-resourced languages, however, although the number of annotated corpora is increasing, the dearth of material continues to make applying these approaches difficult. Our aim is to improve this situation for Basque by both improving coreference resolution and facilitating the creation of a larger corpus for future work on similar tasks.

We will design a semi-automatic hybrid system to speed up corpus tagging by facilitating human annotation. Our system will allow the annotation tool to

[1] http://conll.cemantix.org/2011/task-description.html

display nominal and pronominal coreference chains in a user-friendly and easy-to-understand manner, so that coreferences can be tagged or corrected with a simple click of the mouse.

We will annotate, at coreference level, the Reference Corpus for the Processing of Basque (EPEC), a 300,000 word collection of written standard Basque that has been automatically tagged at different levels (morphology, surface syntax, and phrases). We will endeavor to solve nominal coreferences by combining rule-based techniques and machine learning approaches to pronominal anaphora resolution. The machine learning techniques will allow us to make use of existing resources for Basque, while using rule-based techniques for nominal coreference resolution will allow us to partially circumvent the problem of the limits of those resources.

Our machine learner is trained on a part of the Eus3LB Corpus[2] [15], a collection of previously parsed journalistic texts. This corpus, the basis of the EPEC corpus, has been manually tagged at coreference level, but only pronominal anaphora are annotated [1].

The paper is organized as follows. Section 2 describes some of the most significant work on coreferences, followed by a section presenting the tools we use to annotate corpora automatically and our aim in carrying out this work. In section 4 we describe the automatic coreference resolution process, which is divided into two parts: nominal coreference resolution process and pronominal anaphora resolution process. Section 5 then presents our experimental results and section 6 closes the paper with some concluding remarks.

2 Related Work

Recent work on coreference resolution has been largely dominated by machine learning approaches. In the SemEval-2010 task on Coreference Resolution in Multiple Languages[3] [24], most of the systems were based on those techniques [7, 26, 12]. Nevertheless, rule-based systems have also been applied successfully: in the CoNLL-2011 Shared Task, for example, the best result was obtained by [13], which proposes a coreference resolution system that is an incremental extension of the multi-pass sieve system proposed in [23]. This system is shifting from the supervised learning setting to an unsupervised setting.

At the same time, most other systems proposed at CoNLL-2011 [8, 6, 10] were based on machine learning techniques. The advantage of these approaches is that there are many open-source platforms for machine learning and machine learning based coreference such as BART[4] [27] or the Illinois Coreference Package [5].

The state of the art for languages other than English varies considerably. A rule-based system for anaphora resolution in Czech is proposed in [14], which uses Treebank data containing more than 45,000 coreference links in almost 50,000

[2]Eus3LB is part of the 3LB project.

[3]`http://stel.ub.edu/semeval2010-coref/systems`

[4]`http://www.bart-coref.org/`

manually annotated Czech sentences. Most recently, a substantial amount of newly annotated data for Czech has prompted the application of a supervised machine learning approach to resolving noun phrase coreferences in Czech [20]. On the other hand, [16] presents an approach to Persian pronoun resolution based on machine learning techniques. Other authors present an end-to-end coreference resolution rule-based system for Polish [21].

3 The Tagging Process

The annotation of coreference in Basque starts out with an annotated corpus that provides us with an easier work environment, one that focuses on the specific structures that could be part of a reference chain. The EPEC corpus has been morphosyntactically analyzed by means of MORFEUS [2]. After that, two automatic taggers (rule-based and stochastic) disambiguate at the lemmatization level. Finally, entities, chunks and complex postpositions are identified by means of the following tools: i) EIHERA, which identifies entities (Institution, Person and Location) [3]; and ii) IXATI Chunker [11], which identifies verb chains, noun phrase units, and complex postpositions.

Referring to the annotation of pronominal anaphora, 25.000 words of this corpus was carried out manually. For this tagging, we used the MMAX2 application [19] (adapted to the established requirements). Although the annotation tool is adequate, the process is still arduous and time consuming; we wanted to make it faster and more user-friendly. Toward this end, we developed an automatic coreference resolution system and transported the results it produced into the MMAX2 output window. Thus, depending on the type of coreference, the tool now displays either the possible chains or the possible antecedents. Coreference mentions appear highlighted in the text, so that simply by clicking on a coreference the annotator can see the chain of elements belonging to the same cluster. For each pronominal anaphor, the five most probable antecedents are linked, and the most probable is highlighted. The annotator needs only to choose the correct one.

4 The Coreference Resolution Process

The input of the coreference resolution module consists of a part of the EPEC corpus where each word of the corpus contains its form, lemma, category, POS and morphosyntactic features such as case and number. In addition, NPs are also labeled in the corpus. We only take into account NPs as potential mentions to be included in a coreference chain. The boundaries of these NPs are defined using three labels (*NP*, *NPB*, and *NPE*): if the NP contains more than a word, one label indicates the beginning [NPB] and the other one indicates the end [NPE]. Otherwise, if the NP contains only one word, the word is tagged with a unique label [NP] at its end.

Correct noun phrase tagging is crucial to coreference resolution: if a noun phrase is tagged incorrectly in the corpus, a potential anaphor or antecedent will be lost. We have detected 124 mislabeled noun phrases in our corpus, representing 9% of the total number of NPs. Most of these cases have a beginning label but no end label, an error that is due to the use of Constraint Grammar to annotate NPs automatically. This formalism does not verify if the beginning label has been annotated when it annotates an end label and vice versa. To avoid having this problem hamper our coreference task, we attempted to correct some of these mislabeled cases some simple rules, with varying success. These rules look for the opening and the ending label. Therefore, if one of them lacks, the heuristic tags the missing one. After the corrections were carried out, we started the coreference resolution process with 1265 correct noun phrases.

We divided our coreference resolution process into two subtasks depending on the type of coreference: nominal coreference resolution and pronominal coreference resolution. We used a rule-based method to solve nominal coreferences while employing a machine learning approach to find pronominal anaphora and their antecedents.

4.1 Nominal Coreference Resolution

We implemented the nominal coreference resolution process as a succession of three steps: (1) classifying noun phrases in different groups depending on their morphosyntactic features; (2) searching and linking coreferences between particular groups, thus creating possible coreference clusters; and (3) attempting to eliminate incorrect clusters and return correct ones by means of the coreference selector module, which takes into account the order of the noun phrases in the text.

4.1.1 Classification of Noun Phrases

In order to find easily simple coreference pairs, we classify noun phrases into seven different groups (G1...G7) according to their morphosyntactic features. Some of these features are proposed in [28]. To make an accurate classification, we create extra groups for the genitive constructions.

G1: The heads of those NPs that do **not** contain any **named entity** (see [23]). Although we use the head of the noun phrase to detect coreferences, the entire noun phrase has been taken into account for pairing or clustering purposes.

For example: *Escuderok [euskal **musika** tradizionala] eraberritu eta indartu zuen*. (Escudero renewed and gave prominence to [traditional Basque **music**]). The head of this NP is *musika* ("music"). Hence, we save this word in the group.

G2: NPs that contain **named entities with genitive form**. For example: *[**James Bonden** lehen autoa] Aston Martin DB5 izan zen*. ([**James Bond's** first car] was an Aston Martin DB5).

G3: NPs that contain **named entities with place genitive form**. For example: *[**Bilboko**] biztanleak birziklatze kontuekin oso konprometituak daude*. ([The citizens **of Bilbao**] are very involved in recycling).

G4: NPs that contain **name entities with place genitive form and genitive form** (in Basque, the *-ko* and *-en* suffixes). In other words, this group contains NPs that fulfill the conditions for both G2 and G3. For example:

*[**Jesulinen Bilboko** etxea] Bilboko lekurik onenean dago.* ([**Jesulin's Bilbao** house] is located in the best area of Bilbao).

G5: NPs that contain **named entities**. These named entities can not have any form of genitive or place genitive. For example: *[**Leo Messi**] munduko futbol jokalaririk hoberena izango da segur aski.* ([**Leo Messi**] is probably the best football player in the world).

G6: Appositions + named entities. For example: *[**Pakito Mujika presoa**] orain dela gutxi epaitu dute.* ([**The prisoner Pakito Mujika**] has been judged recently).

G7: Postpositional phrases. Basque has a postpositional system (instead of prepositions, as in English), and therefore we mark the independent postposition and the preceding NP as a unit. For example: *Joan den astean [**Moriren aurka**] aurkeztutako zentsura mozioak krisia sortu zuen LDPn.* (The vote of no confidence [**against Mori**] caused a crisis in the LDP last week).

4.1.2 Candidate Selection

The aim of this module is to find the most obvious clusters of coreferences that will then be evaluated in the next module. To create these clusters we consider two main factors: (1) how to match mentions to decide if they are relevant candidates and (2) in which groups we must search for candidate mentions.

To decide whether two particular mentions are coreferential, we use two different matching techniques. Which one we select depends on the number of words in the mentions: if the number of words is the same in both cases, we use Exact Match, otherwise we use Relaxed Head Matching [23]. Let us explain these two mechanisms.

Exact Match (EM): To consider two mentions with the same number of words coreferential, they have to be equal character for character. In the case of the mentions of the no named entities group (G1), we use the mention heads for matching. For example: *[The **dog**] was huge and running free... [This **dog**] is the same one that bit John.*

Those two noun phrases belong to group 1 and their heads (dog) coincide character for character. So for the time being we consider them potential coreferences and save them, pending a final decision in the last module.

Relaxed Head Matching (RHM): We consider two mentions potential coreferences when all the words of the shortest mention are included in the longest mention with no alteration of order and with no intermediate words. For example, the system matches the mention *James Morrison* to the mention *James Morrison Strous* because all the words that compose the shortest mention are in the other mention without any alteration. But it does not match the mention *James Morri-*

	G1	G2	G3	G4	G5	G6	G7
G1	x	x	x	x		x	
G2	x	x					
G3	x		x				
G4	x			x			
G5					x	x	
G6	x				x	x	
G7							x

Table 1: Possible combinations

son to the mention *James Lewis Morrison*, because although all the words of the first mention are included in the second mention, the second mention contains an intermediate word (Lewis).

In order to decide in which groups to look for candidate mentions, we can use our knowledge of the composition of the groups. Thus, we know that coreferences between some groups are more likely than between others. For example, we can expect to find coreferences between the mentions of group 5 and group 6 because both groups include named entities, like the mentions *Brad Pitt* (G5) and *the actor Brad Pitt* (G6). By contrast, there will be no coreferences between the elements of the group 1 and group 5, for the simple reason that group 1 is created with no named entities and group 5 is created with named entities only. Consider the mention *carpenter* (G1) and the mention *John Carpenter* (G5). Although the word *carpenter* appears in both mentions, in the first one it refers to a profession and in the second to a surname. Therefore, we have discarded some combinations; the possible ones are summarized in Table 1.

Once each mention is filed in its proper group and it is clear what combinations of groups we need to look for possible coreferences in, we can begin searching for possible coreferences for each mention.

First, we search the mentions of each group for possible candidates. Consider a group with the mentions *Juan Montero, Perez, Montero, Gallastegi, Buesa*. The system would link the mentions *Juan Montero* and *Montero*. For some groups (G1 and G5), the system next tries to link mentions with mentions in other groups. For example, the mentions included in group 1 [*the dog*...] are the most general, so we match them with mentions from several groups (G2 [*George's dog*...], G3 [*the dog in Rome*...], G4 [*Nuria's Madrid dog*...], G6 [*the dog Larry*...]) due to the probability of finding coreferences in these morphologically compatible groups. However, it is also possible that two mentions in different groups could be coreferential only through a third mention. For example, we cannot directly join the mentions *Michael Jordan* and *basketball player* because we lack a clue that we could use to make the connection. But if we were to find the mention *Michael Jordan the basketball player*, we could use it to join all three mentions.

4.1.3 Coreference Selector Module

The objective of this module is to return clusters of coreferences, validating or modifying the clusters that it receives from the previous module. The input of this module, then, is a set of clusters that links coreference candidates.

In order to decide whether a cluster of coreferences is correct, the order in which the mentions of the cluster appear in the text is crucial. We can find the same mention in the text twice without it being coreferential. Consider this example: *[Iñaki Perurena] has lifted a 325-kilo stone... [Perurena] has trained hard to get this record... His son [Jon Perurena] has his father's strength... [Perurena] has lifted a 300-kilo cube-formed stone*. In this example we find the mention Perurena twice. The previous module links these mentions, creating a cluster of four mentions [*Iñaki Perurena, Perurena, Jon Perurena, Perurena*]. However, *Jon Perurena* and *Iñaki Perurena* are not coreferential, so this cluster is not valid. To eliminate this type of erroneous linkage, the coreference selector module takes into account the order in which the mentions appear. In other words, it matches the mention *Iñaki Perurena* to the first *Perurena* mention and the mention *Jon Perurena* to the second *Perurena* mention, as this is most likely to have been the writer's intention.

The system labels all marked mentions as possible coreferents (m_1, m_2, m_3, m_4, m_n) and then proceeds through them one by one trying to find coreferences. The system uses the following procedure. (1) If there exists a mention (for example m_3) that agrees with an earlier mention (for example m_1) and there **does not exist** a mention between them (for example m_2) that is coreferential with the current mention (m_3) and not with the earlier mention (m_1), the module considers them (m_1 and m_3) coreferential. (2) If there **exists** a mention between the two mentions (for example m_2) that is coreferential with the current mention (m_3) but not with the earlier mention (m_1), the module will mark as coreferents the intermediate mention (m_2) and the current mention (m_3). Thus, the module forms, step by step, different clusters of coreferences, and the set of those strings forms the final result of the Nominal Coreference Resolution Process.

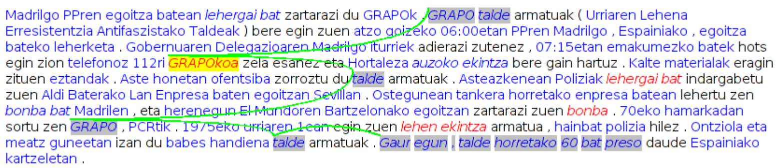

Figure 1: Nominal coreference example.

Figure 1 shows the result of the Nominal Coreference Resolution represented by the MMAX2 tool. The annotator then checks if the cluster of coreferences is correct. If it is, the coreference is annotated simply by clicking on it; if it is not, the annotator can easily correct it. Thus, the time employed in annotating coreferences is reduced.

4.2 Pronominal Anaphora Resolution

In order to use a machine learning method, a suitable annotated corpus is needed. As noted in the introduction, we use part of the Eus3LB Corpus. This corpus contains 349 annotated pronominal anaphora and it's different from the data we use to evaluate and develop our pronominal and nominal coreference resolution systems. The method used to create training instances is similar to the one explained in [25]. Positive instances are created for each annotated anaphor and its antecedents, while negative instances are created by pairing each annotated anaphor with each of the preceding noun phrases that are between the anaphor and the antecedent. Altogether, we have 968 instances; 349 of them are positive, and the rest (619) negative.

The method we use to create testing instances is the same we use to create training instances—with one exception—. As we can not know *a priori* what the antecedent of each pronoun is, we create instances for each possible anaphor (pronouns) and the eight noun phrases nearest to it. We choose the eight nearest NPs because experiments on our training set revealed that the antecedent lies at this distance 97% of the time. Therefore, for each anaphor we have eight candidate antecedents, i.e., eight instances. The features used are obtained from the linguistic processing system defined in [4].

The next step is to use a machine learning approach to determine the most probable antecedents for each anaphor. After consultation with linguists, we decided on returning a ranking of the five most probable antecedents. The most probable antecedent will be highlighted. Then, these clusters of coreferences are displayed in the MMAX2 tool. In this manner, the annotator will select the correct antecedent for each anaphor from a set of five possible antecedents, instead of having to find it in the whole text, saving time and improving performance. Consequently, we will be able to create a larger tagged corpus faster as well as achieve better models for applying and improving the machine learning approach.

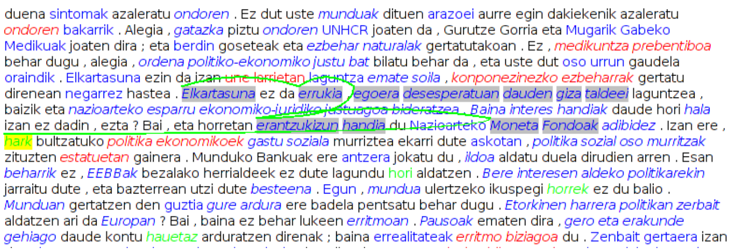

Figure 2: Pronominal coreference example.

Figure 2 represents an anaphor (*hark*, in English *he/she/it*) and its five most probable antecedents in the MMAX2 window. The annotator can choose the correct antecedent of the pronominal anaphor with a few clicks.

Nominal	P	R	F1
MUC	75.33%	81.33%	78.21%
B^3	72.80%	83.95%	77.97%
BLANC	90.5%	87.57%	88.98%
Pronominal	76.9%	60.0%	67.4%

Table 2: Results of the anaphora resolution system.

5 Experimental Results

We use two different strategies to evaluate the two coreference resolution processes, since we use two different methods to link coreferences. In the pronominal anaphora resolution process, we return the five most probable antecedents for each anaphor, while in the nominal coreference resolution process we return a cluster of mentions that links coreferential mentions for each nominal coreference.

The evaluation metrics are chosen with a view toward appropriateness. We use the classic measures (precision, recall and F1) to evaluate the pronominal anaphora resolution process, counting as success the instances when the real antecedent of the pronominal anaphor is among the five most probable antecedents. To evaluate nominal coreferences, we use BLANC, MUC and B^3 metrics, as they are the three most significant metrics used for this task.

We use 1004 NPs to develop our nominal coreference resolution system and 281 NPs to evaluate it. For the evaluation of our pronominal anaphora resolution system, we use 130 pronominal anaphora of those 1285 NPs.

We present the results of our coreference resolution system in Table 2. In the nominal coreference resolution system we obtain an F-score of at least 78% using the three above-mentioned metrics. On the other hand, using the pronominal coreference resolution system, the F-score is 67.4%. Although these results are not the best obtained in coreference resolution systems, they build a solid base for improving our system and indicate that our system is of considerable use in speeding up the manual nominal/pronominal anaphora annotation. This, in turn, will allow us to create a broader corpus and use it to improve our hybrid approach to automatic corpus annotation.

6 Conclusions and Future Work

In this work we present a system for automatically annotating nominal and pronominal coreferences using a combination of rules and ML methods. Our work begins by detecting incorrectly tagged NPs and, in most cases, correcting them, recovering 63% of the incorrectly tagged NPs. Next, in the case of the nominal coreferences, we divide the NPs into different groups according to their morphological features to find coreferences among the compatible groups. Then we use a ML approach to solve pronominal anaphora; this returns, for each anaphor, a cluster that contains the anaphor and its five most probable antecedents.

Our results demonstrate that, despite their simplicity, ML and deterministic models for coreference resolution obtain competitive results. This will allow us to create an automatic annotation system to improve the manual annotation process of the corpora. A larger tagged corpus, in turn, will enable us to improve the performance of our automatic system.

Acknowledgements

This work has been supported by Hibrido Sint (TIN2010-20218) project.

References

[1] I. Aduriz, K. Ceberio, and A. Díaz de Ilarraza. Pronominal anaphora in Basque: Annotation issues for later computational treatment. In *DAARC2007 Lagos (Portugal); pp. 1-7, ISBN: 978-989-95343-0-8*, 2007.

[2] I. Alegria, X. Artola, K. Sarasola, and M. Urkia. Automatic morphological analysis of Basque. 1996.

[3] I. Alegria, N. Ezeiza, and I. Fernandez. Named entities translation based on comparable corpora. In *Multi-Word-Expressions in a Multilingual Context Workshop on EACL06. April 3. pp.1-8. Trento*, 2006.

[4] O. Arregi, K. Ceberio, A. Díaz de Ilarraza, I. Goenaga, B. Sierra, and A Zelaia. A first machine learning approach to pronominal anaphora resolution in Basque. In *IBERAMIA 2010. LNAI 6433*, 2010.

[5] E. Bengtson and D. Roth. Understanding the value of features for coreference resolution. In *In EMNLP*, 2008.

[6] A. Björkelund and P. Nugues. Exploring lexicalized features for coreference resolution. In *Proceedings of the Fifteenth Conference on Computational Natural Language Learning: Shared Task*, CONLL Shared Task'11, pages 45–50, Stroudsburg, PA, USA, 2011.

[7] S. Broscheit, M. Poesio, S. Paolo Ponzetto, K. J. Rodriguez, L. Romano, O. Uryupina, Y. Versley, and R. Zanoli. BART: A multilingual anaphora resolution system. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, SemEval'10, pages 104–107, Stroudsburg, PA, USA, 2010.

[8] K. Chang, R. Samdani, A. Rozovskaya, N. Rizzolo, M. Sammons, and D. Roth. Inference protocols for coreference resolution. In *Proceedings of the Fifteenth Conference on Computational Natural Language Learning: Shared Task*, CONLL Shared Task '11, pages 40–44, Stroudsburg, PA, USA, 2011.

[9] G. Doddington, A. Mitchell, M. Przybocki, L. Ramshaw, S. Strassel, and R. Weischedel. The Automatic Content Extraction (ACE) Program–Tasks, Data, and Evaluation. *Proceedings of LREC 2004*, pages 837–840, 2004.

[10] C. N. dos Santos and D. L. Carvalho. Rule and tree ensembles for unrestricted coreference resolution. In *Proceedings of the Fifteenth Conference on Computational Natural Language Learning: Shared Task*, CONLL Shared Task'11, pages 51–55, Stroudsburg, PA, USA, 2011.

[11] J. M. Arriola A. Atutxa A. Diaz-De-Ilarraza N. Ezeiza K. Gojenola M. Oronoz A. Soroa R. Urizar I. Aduriz, M. J. Aranzabe. Methodology and steps towards the construction of EPEC, a corpus of written Basque tagged at morphological and syntactic levels for the automatic processing. pages 1–15. Rodopi. Book series: Language and Computers, 2006.

[12] H. Kobdani and H. Schütze. Sucre: A modular system for coreference resolution. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, SemEval'10, pages 92–95, Stroudsburg, PA, USA, 2010.

[13] H. Lee, Y. Peirsman, A. Chang, N. Chambers, M. Surdeanu, and D. Jurafsky. Stanford's Multi-Pass Sieve Coreference Resolution System at the CoNLL-2011 Shared Task. In *Proceedings of the CoNLL-2011 Shared Task*, 2011.

[14] N. G. Linh. Rule-based approach to pronominal anaphora resolution applied on the Prague dependency treebank 2.0 data, 2007.

[15] A. Díaz de Ilarraza L. Moreno E. Bisbal-M. J. Aranzabe A. Ageno M. A. Marti F. Navarro M. S. Palomar, M. Civit. Construcción de una base de datos de arboles sintáctico-semánticos para el catalán, euskera y español. In *XX. Congreso SEPLN*, 2004.

[16] N. S. Moosavi and G. Ghassem-Sani. A ranking approach to Persian pronoun resolution. Advances in Computational Linguistics. In *Advances in Computational Linguistics. Research in Computing Science 41*, pages 169–180, 2009.

[17] MUC-6. Coreference task definition (v2.3, 8 sep 95). In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 335–344, Columbia, Maryland, USA, November 6–8 1995.

[18] MUC-7. Coreference task definition (v3.0, 13 jul 97). In *Proceedings of the 7th Message Understanding Conference (MUC-7)*, Fairfax, Virginia, USA, April 29–May 1 1998.

[19] C. Müller and M. Strube. Multi-level annotation of linguistic data with MMAX2. In Sabine Braun, Kurt Kohn, and Joybrato Mukherjee, editors, *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, pages 197–214. Peter Lang, Frankfurt a.M., Germany, 2006.

[20] M. Novák and Z. Žabokrtský. Resolving noun phrase coreference in Czech. In *Proceedings of the 8th international conference on Anaphora Processing and Applications*, DAARC'11, pages 24–34, Berlin, Heidelberg, 2011. Springer-Verlag.

[21] M. Ogrodniczuk and M. Kopeć. End-to-end coreference resolution baseline system for Polish, 2011.

[22] M. Poesio. The mate/gnome proposals for anaphoric annotation, revisited. In *In Michael Strube and Candy Sidner (editors), Proceedings of the 5th SIGdial Workshop on Discourse and Dialogue*, pages 154–162, 2004.

[23] K. Raghunathan, H. Lee, S. Rangarajan, N. Chambers, M. Surdeanu, D. Jurafsky, and C. D. Manning. A multi-pass sieve for coreference resolution. In *EMNLP*, pages 492–501, 2010.

[24] M. Recasens, L. Màrquez, E. Sapena, M. A. Martí, M. Taulé, V. Hoste, M. Poesio, and Y. Versley. Semeval-2010 task 1: Coreference resolution in multiple languages. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, SemEval'10, pages 1–8, Stroudsburg, PA, USA, 2010.

[25] W. M. Soon, D. Chung, Y. Lim, and H. T. Ng. A machine learning approach to coreference resolution of noun phrases, 2001.

[26] O. Uryupina. Corry: A system for coreference resolution. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, SemEval'10, pages 100–103, Stroudsburg, PA, USA, 2010.

[27] Y. Versley, S. P. Ponzetto, M. Poesio, V. Eidelman, A. Jern, J. Smith, X. Yang, and A. Moschitti. Bart: a modular toolkit for coreference resolution. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Demo Session*, HLT-Demonstrations '08, pages 9–12, Stroudsburg, PA, USA, 2008. Association for Computational Linguistics.

[28] R. Vieira and M. Poesio. An empirically based system for processing definite descriptions. *Comput. Linguist.*, 26(4):539–593, December 2000.