FACHLITERATUR
EDITION ROMIOSINI
ΕΠΙΣΤΗΜΗ

ICGL12 | 12th INTERNATIONAL CONFERENC ON GREEK LINGUISTICS
16 – 19 SEPTEMBER 2015
FREIE UNIVERSITÄT BERLIN, CEMOG

Proceedings of the ICGL12

vol. 2

The International Conference on Greek Linguistics is a biennial meeting on the study and analysis of Greek (Ancient, Medieval and Modern), placing particular emphasis on the later stages of the language.

PROCEEDINGS OF THE ICGL12
ΠΡΑΚΤΙΚΑ ΤΟΥ ICGL12

Thanasis Georgakopoulos, Theodossia-Soula Pavlidou, Miltos Pechlivanos, Artemis Alexiadou, Jannis Androutsopoulos, Alexis Kalokairinos, Stavros Skopeteas, Katerina Stathi (Eds.)

PROCEEDINGS OF THE 12TH INTERNATIONAL CONFERENCE ON GREEK LINGUISTICS

ΠΡΑΚΤΙΚΑ ΤΟΥ 12ΟΥ ΣΥΝΕΔΡΙΟΥ ΕΛΛΗΝΙΚΗΣ ΓΛΩΣΣΟΛΟΓΙΑΣ

VOL. 2

Vertrieb und Gesamtherstellung: Epubli (www.epubli.de)
Satz und Layout: Rea Papamichail / Center für Digitale Systeme, Freie Universität Berlin
Gesetzt aus Minion Pro
Umschlaggestaltung: Thanasis Georgiou, Yorgos Konstantinou
Umschlagillustration: Yorgos Konstantinou

ISBN 978-3-946142-35-5
Printed in Germany

Online-Bibliothek der Edition Romiosini:
www.edition-romiosini.de

ΠΕΡΙΕΧΟΜΕΝΑ

CO-REFERRING OR NOT CO-REFERRING? ANSWER THE QUESTION!

Marina Vassiliou[1], Angelos Georgaras[2],Prokopis Prokopidis[1] & Haris Papageorgiou[1]
[1]Institute for Language and Speech Processing, [2]University of Athens
{mvas, prokopis, xaris}@ilsp.gr, angelosgeorgaras@hotmail.com

Περίληψη

Επίλυση συναναφοράς είναι ο εντοπισμός εκφράσεων που αναφέρονται στην ίδια οντότητα. Βασική πτυχή στην Επεξεργασία Φυσικής Γλώσσας, συμβάλλει σε εφαρμογές όπως εξόρυξη πληροφορίας, περίληψη κειμένου και συστήματα ερωταποκρίσεων. Στο παρόν άρθρο περιγράφεται ένα σύστημα επίλυσης συναναφοράς, το πρώτο, εξ όσων γνωρίζουμε, για την Ελληνική. Χειρίζεται όλες τις εκφράσεις (ονοματικές και αντωνυμικές, κενά υποκείμενα) και βασίζεται σε ένα κανονιστικό αλγόριθμο, ο οποίος, εφαρμόζοντας μία σειρά – κυρίως μορφολογικών και συντακτικών – κριτηρίων δημιουργεί συστάδες συναναφερόμενων εκφράσεων. Η πρώτη υλοποίηση του συστήματος επιτυγχάνει πολύ ικανοποιητικά ποσοστά ακρίβειας (~ 83%). Επίσης αναπτύχθηκε σχήμα επισημείωσης σχέσεων συναναφοράς που εφαρμόστηκε σε ένα σώμα κειμένων μεγέθους >90.000 λέξεων.

Keywords: coreference, near identity, bridging, coreference resolution, coreference annotation, Natural Language Processing

1. Introduction

Coreference Resolution (CR) denotes the detection of all linguistic expressions in a text which refer to the same discourse entity. It has been (and still is) one of the most intriguing tasks within Natural Language Processing over the past twenty years (see

Ng (2010) for a concise review of the field). An interesting problem per se, coreference resolution also assumes an ancillary role in other NLP tasks such as information extraction, text summarisation, machine translation and question answering.

Coreference resolution is construed as a clustering problem, i.e. the coreferent expressions within a text must be linked as members of the same chain and grouped into the same cluster. The solution to this clustering problem has been attempted by various systems, which are either based on machine learning algorithms (Soon et al., 2001, Ng and Cardie, 2002, Versley et al., 2008b, Durrett and Klein, 2013) or follow rule-based approaches (Lappin and Leass, 1994, Haghighi and Klein, 2009, Lee et al., 2013, O'Connor and Heilman, 2013). Machine learning systems, supervised and unsupervised ones, have taken the lead over the last decade due to the availability of manually annotated corpora; yet rule-based systems still outperform them in most cases.

Models for handling coreference could be classified into three types: First, the mention[1]-pair models, where expressions in a text are examined in pairs and then local decisions are assembled together (Soon et al., 2001, Ng and Cardie 2002a, Versley et al., 2008b). Alternatively, the entity-mention models highlight the importance of simultaneously checking coreference over all mentions in text which refer to an entity (Luo et al., 2004, Yang et al., 2008a, 2004, Haghighi and Klein, 2010, Lee et al., 2011). Finally, ranking models arrange all candidate antecedents of an expression based on probability (Iida et al., 2003, Yang et al., 2008b).

CR systems utilise a wide range of features, varying between simple string matching and grammatical features (part of speech, gender, number etc.) to syntactic and semantic features (e.g. agreement, binding constraints, syntactic role, selectional restrictions, θ-roles, semantic relations etc.). Furthermore, in order to enhance precision, there have been attempts to improve coreference resolution by feeding systems with knowledge of the world, the rationale being that humans also resort to this knowledge in order to process and understand texts. This knowledge is drawn from web resources such as Wikipedia, DBPedia or YAGO (cf. Uryupina 2006, Ponzetto et al., 2007, Versley et al., 2008a, Bryl et al., 2010).

In this paper we describe a coreference resolution system, the first one, to our knowledge, that handles Greek data. The system implements a rule-based algorithm, which falls under the category of the entity-mention models, and targets both precision and recall. More specifically, with the aim of maximising recall, it initially extracts all lin-

1 *Mention* denotes the occurrence of an expression in a text.

guistic expressions from a text and creates corresponding clusters (i.e. coreference chains). Subsequently, a series of precision-oriented criteria, which primarily rely on morphological and syntactic features, are applied in a stepwise mode, checking the consistency of already formed clusters and governing the creation of new ones.

The linguistic expressions handled by the system comprise only nominal and pronominal referential expressions as well as null subjects in both finite and gerundive clauses. Adverbial expressions with a referential function are not currently supported.

Furthermore, the current article reports on related resources that have been created in parallel, namely a coreference annotation scheme for the Greek language that also addresses bridging and near identity relations, as well as a Greek corpus annotated with these relations.

The paper is structured as follows: Section 2 provides information on the coreference resources for the Greek language, namely the annotation scheme and the corresponding corpus. Section 3 describes the coreference resolution system, while Section 4 reports on the evaluation results obtained. The concluding section outlines future directions with respect to the development of the system and the improvement of its performance.

2. Coreference annotation scheme for the Greek language

2.1 Defining relations between expressions

The Greek coreference annotation scheme specifies three types of relations between linguistic expressions, (i) coreference, (ii) near identity and (iii) bridging.

Coreference denotes a relation between (at least) two linguistic expressions, which have the same referent, i.e. they point to the same discourse entity. The linking of these expressions results in a chain, either within the sentence boundaries (examples 1-3) or spanning more than one sentence (example 4):

(1) Η εξέλιξη αυτή ανακοινώθηκε ταυτόχρονα στις ΗΠΑ και τη Βόρεια **Κορέα** και προέκυψε δύο μήνες μετά το θάνατο του ηγέτη της Βόρειας **Κορέας**.

'This development was announced in the US and North Korea at the same time and came two months after the death of the leader of North Korea.'

(2) Στην Ουάσιγκτον, ο Λευκός Οίκος εξέφρασε ικανοποίηση για την απόφαση της **Πιονγκγιάνγκ**, χαρακτηρίζοντάς **την** θετικό πρώτο βήμα.

'In Washington, the White House welcomed the decision of Pyongyang and characterised it as a positive first step.'

(3) Πέρασε ο ίδιος νύχτα στο **Ακρωτήρι όπου** ύψωσε την ελληνική σημαία.

'He went himself to Akrotiri during the night and raised the Greek flag.'

(4) Ο Αμερικανός **Πρόεδρος** δήλωσε … ότι "προτιμά το θέμα να επιλυθεί μέσω της διπλωματίας." Ο Μπαράκ **Ομπάμα** ανέφερε ότι «ήδη υπάρχουν πολλές αναφορές για πόλεμο με το Ιράν."

'The US president said … that "the issue will preferably be resolved through diplomacy." Barack Obama said that "there are already many references to war with Iran".'

Near identity denotes a quasi coreferential relation between two different linguistic expressions or instances of the same expression, which implicitly refer to the same entity or represent an aspect of the same entity. These expressions "are partially the same in that they share most of the important characteristics, but differ in at least one crucial dimension" (Recasens and Martí, 2010: 151). In example (5) the two instances of the expression *"ΗΠΑ"* are quasi coreferent, since the first one has a locative reading, whereas the second one refers to a related aspect of this location, namely the US government.

(5) Η νίκη του καλωσορίστηκε τόσο στην Κούβα όσο και στις **ΗΠΑ**. … Έτσι οι **ΗΠΑ** έγιναν πολύ εχθρικές έναντι του Κάστρο κατά το 1959.

'His victory was welcomed in Cuba as well as in the United States … Therefore, in 1959 the United States became pretty hostile towards Castro.'

Bridging denotes an associative relation between two non-coreferent linguistic expressions. This relation may come in various types (for instance set-subset, part-whole, entity-property, contrastive association etc.). Example (6) instantiates a set-subset relation holding between the two bold-faced expressions.

(6) Σε όλες σχεδόν τις κατοικημένες **περιοχές** του **κέντρου** υπάρχει η ίδια εικόνα ερήμωσης.

'One sees the same picture of desolation in almost all populated downtown areas.'

2.2 The annotation scheme in a nutshell

The annotation scheme is largely based on the PDT 2.0 (Prague Dependency Treebank)[2] scheme, used for manual annotation of texts at multiple levels (morphology, syntax, semantics and pragmatics); yet it contains several modifications pertinent to the Greek language, while it was also extended to accommodate near identity relations.

All kinds of expressions are taken into account (nominal expressions and pronominal elements, anaphoric adverbs, and null subjects). Notably, as regards coreference, split antecedents are allowed. Additionally, metonymy, which functions as a coreference mechanism, is also covered in the annotation scheme. Furthermore, near identity relations come in five types. Finally, nine bridging relation types are specified.

A comprehensive description of the scheme together with annotation guidelines is available online[3].

2.3 The annotated corpus

The resource annotated with the aforementioned scheme is a medium-sized corpus comprising mainly news texts drawn from various sources. Table 1 on the next page illustrates the corpus profile.

Figure 1 illustrates the TrEd[4] annotation tool that was used for the annotation of the corpus. Text is represented as sequences of dependency trees and the coreferent nodes are linked with an arrow.

2 https://ufal.mff.cuni.cz/pdt2.0/

3 http://gdt.ilsp.gr/guidelines/coreference/coreferenceannotationguidelines.pdf/@@download/file/coreferenceannotationguidelines.pdf

4 http://ufal.mff.cuni.cz/tred/

Source	Texts	Sentences	Tokens
newspapers	11	186	4.120
elwikinews	92	675	14.739
wikipedia	11	263	6.263
europarl	45	1.301	30.076
setimes	6	67	1.684
voa	9	89	2.551
in.gr	81	791	21.843
opinion texts	20	591	11.788
Total	**275**	**3.963**	**93.064**

Table 1 | Coreference-annotated corpus

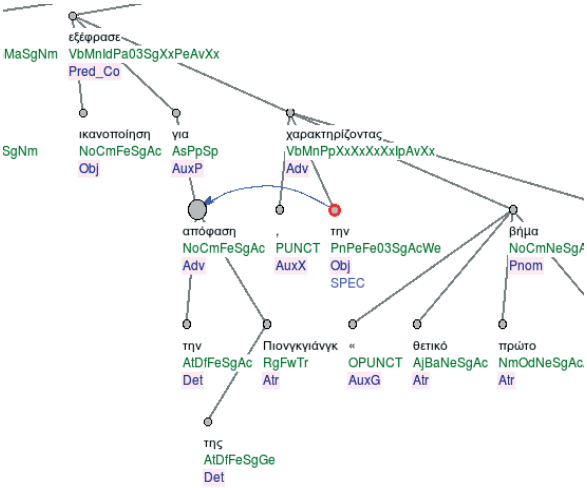

Figure 1 | Screenshot of the annotation environment

3. A coreference resolution system for Greek

Our coreference resolution system for Greek belongs to the rule-based paradigm rather than the machine learning one. This decision was dictated not only by the scarcity of available training resources in Greek, but rather by the fact that rule-based approaches yield the best results as yet. Following the entity-mention model of dealing with coreference, the system attempts to jointly detect all the expressions in a text, nominal and

pronominal ones, phonologically instantiated or not, which refer to the same entity.

The system receives as input an XML file (see Figure 2) containing text annotated with information about token, lemma, Part-of-Speech (PoS) tag and syntactic dependency relations.

```
<children id="t-20140701-ingr-8013-s1w7">
<t_lemma>Ρωσία</t_lemma>
<deepord>7</deepord>
   <children id="t-20140701-ingr-8013-s1w6">
   <t_lemma>ο</t_lemma>
   <deepord>6</deepord>
   <form>της</form>
   <tag>AtDfFeSgGe</tag>
   <sentord>6</sentord>
   <afun>Det</afun>
   </children>
<form>Ρωσίας</form>
<tag>NoPrFeSgGe</tag>
<sentord>7</sentord>
<afun>Atr</afun>
</children>
```

Figure 2 | *Indicative input to the CR system*

s1w6	ο	της	AtDfFeSgGe	Det	
s1w7	**Ρωσία**	**Ρωσίας**	NoPrFe03SgGe	Atr	**(7)**
s1w8	λόγω	λόγω	AsPpSp	AuxP	
s1w9	ο	των	AtDfFePlGe	Det	
s1w10	**κύρωση**	**κυρώσεων**	NoCmFe03PlGe	Atr	**(4)**
s1w11	.	.	PTERM_P	AuxK	
s2w1	ο	Οι	AtDfFePlNm	Det	
s2w2	**κύρωση**	**κυρώσεις**	NoCmFe03PlNm	Sb	**(4)**
s2w3	**που**	**που**	PnReFe03PlNmXx	Sb	**(4)**
s2w4	έχω	έχουν	VbMnIdPr03PlXxIpAvXx	AuxV	
s2w5	επιβάλλω	επιβληθεί	VbMnNfXxXxXxXxPePvXx	Atr	
s2w6	στου	στη	AsPpPaFeSgAc	AuxP	
s2w7	**Ρωσία**	**Ρωσία**	NoPrFe03SgAc	IObj	**(7)**
s2w23	μπορώ	μπορεί	VbIsIdPr03SgXxIpAvXx	Obj_Co	
s2w24	να	να	PtSj	AuxV	
s2w25	οδηγώ	οδηγήσουν	VbMnIdXx03PlXxPeAvXx	Sb	
s2a1	**#Cor**	-	03PlNm	Sb	**(4)**

Figure 3 | *Indicative output of the CR system*

The system output includes an extra annotation layer whereby each expression is flagged with the id of the cluster it has been classified into. In Figure 3, the expressions marked with the same id (see last column) are grouped into the same cluster, thus they are coreferent.

The CR algorithm is entity-oriented, trying to jointly discover all the expressions that refer to the same entity. Eight phases can be identified:

Phase 1 – Process text: Each word of the input text is represented as a 10-feature vector.

[s6w10, απεργία, NoCmFe03SgGe, 10, Obj, s6w9, 6, 11, 92, απεργίας]

- Feature 1: word id
- Feature 2: word lemma
- Feature 3: PoS tag
- Feature 4: position of the word in the sentence
- Feature 5: dependency relation tag
- Feature 6: parent node id
- Feature 7: sentence id
- Feature 8: clause id
- Feature 9: position of the word in the text
- Feature 10: token

Phase 2 – Identify expressions: All candidate expressions in a text are identified and classified into four categories, namely (i) nominal mentions, (ii) pronominal mentions, (iii) null subjects of a finite clause and (iv) null subjects in non-finite clauses. Expressions which lack anaphoric properties, that is, they do not refer to an entity, for example attributive expressions, are not considered eligible and are excluded.

Phase 3 – Initial clustering: All nominal mentions are clustered based on their lemma. So mentions with the same lemma are considered coreferent and grouped together. Mentions which differ in lemma yet are joined in an appositional structure, for instance *"Ο Βλαντιμίρ* ***Πούτιν****, ο* ***πρόεδρος*** *της Ρωσίας"*, are also considered coreferent and are clustered together.

Phase 4 – Refine clustering: The initial clustering of the nominal mentions is re-examined and modified accordingly on the basis of pre- or post- modifiers. So, mentions that have been grouped together due to lemma identity, will no longer form the same

cluster if their modifiers differ (e.g. *"**αρχική** έκθεση"* vs. *"**τελική** έκθεση"*).

Phase 5 – Handle null subjects: For each null subject the algorithm detects its coreferent mention, which could itself be a null subject, utilising mainly agreement and structural features. If the mention already belongs to a cluster, then the null subject is included too. Otherwise, a two-member cluster is formed comprising the null subject and its antecedent.

Phase 6 – Handle pronominal mentions: The algorithm detects the antecedent of each pronominal element utilising agreement and locality features. For instance, the antecedent of a relative pronoun must precede the pronoun and agree with it in gender and number. In a similar vein to Phase 5, if the antecedent has been already clustered, then it pied-pipes the pronominal mention to its own cluster. In a different case, the pronoun and its antecedent form a two-member cluster.

Phase 7 – Merge clusters: The algorithm re-examines the already-formed clusters and merges the ones that contain coreferent items. For example, the two-member cluster *{Πούτιν, Πρόεδρος}* will be merged with a cluster containing mentions which have the lemma *"πρόεδρος"*.

Phase 8 – Handle residual items: Further merging is attempted concerning smaller clusters which contain null subjects and pronouns. Additionally, singletons, i.e. one-member clusters, are removed.

4. Evaluation of the coreference resolution system

4.1 Evaluation setup

The evaluation of the system was performed on a corpus of texts drawn from the web and the Greek Dependency Treebank[5]. Two subsets (Sets 1 and 2), constituting the largest part of the corpus, belonged to the news domain, while a small corpus subset (Set 3) comprised opinionated texts. The corpus was automatically annotated with information about lemma, PoS tag and dependency relations. The annotations were then manually corrected so as to establish that the CR system's performance could be assessed without being affected by any pre-processing errors, while the remaining errors could be attributed to the algorithm itself.

5 http://gdt.ilsp.gr

For evaluation purposes it is required to compile a golden corpus, i.e. the correct output of the system, against which the system output is compared. To obtain the golden output, a second version of the corpus was created, where the coreferential chains were manually indicated.

Table 2 provides the details of the evaluation corpus.

Sets	Type	Source	Texts	Sentences	Mean number of sentences	Tokens	Mean number of tokens
Set1	news	GDT; web	40	311	7,78	6.913	22,23
Set2	news	in.gr	81	791	9,77	21.843	27,61
Set3	restaurant reviews	athinorama	15	446	29,73	8.711	19,53
			136	**1.548**		**37.467**	

Table 2 | Evaluation corpus

4.2 Evaluation results

A CR system's performance is assessed in terms of its capacity to detect the coreferent expressions within a text and cluster them together. For our experiments a series of established evaluation metrics have been employed, namely MUC (Vilain et al., 1995), B^3 (Bagga and Baldwin, 1998), CEAF (Luo, 2005) and BLANC (Recasens and Hovy, 2011, Luo et al., 2014). These metrics calculate the system's felicity in detecting coreference relations, by comparing the system output to the golden output, and produce recall, precision and *F* scores.

Recall, in this case, measures the number of the correct coreference chains (clusters) that the system succeeds in creating. Precision, on the other hand, denotes how many of the system-created coreference chains (clusters) are correct, while the *F* score indicates the harmonic mean of these two measures. It should be noted that at this point, where the first system version was evaluated, it was necessary to establish that the system was capable of creating correct clusters; subsequently, the precision score was more important than recall.

Table 3 lists the results obtained for the specific dataset. The bold-faced values are the highest ones obtained per set regarding the precision and *F* scores. It is observed that

the average precision for all sets approximates 83%, while the corresponding *F* score is close to 77%.

	Set1			Set2			Set3		
Metrics	**Recall**	**Precision**	***F***	**Recall**	**Precision**	***F***	**Recall**	**Precision**	***F***
MUC	0,78	**0,92**	**0,84**	0,66	**0,83**	**0,73**	0,52	0,72	0,60
B^3	0,74	**0,92**	0,81	0,61	0,82	0,69	0,52	**0,73**	0,60
CEAFM	0,79	0,88	0,83	0,68	0,78	**0,73**	0,59	0,72	0,65
CEAFE	0,83	0,82	0,82	0,75	0,71	0,72	0,68	0,66	**0,67**
BLANC	0,69	0,91	0,77	0,53	0,78	0,61	0,36	0,66	0,45

Table 3 | Evaluation results for the first release of the CR system

From the evaluation results it is evident that the first system release has achieved a promising performance, despite the fact that some phenomena, for instance split antecedents, metonymy or long-distance or inter-sentential dependencies, have by design not been addressed in the first implementation. Therefore, it is expected that there is ground for further improvement.

Furthermore, some preliminary testing has shown that factors such as text size or number of mentions possibly have an impact on the system's performance (see Table 4); yet further experimentation with more texts and various domains is needed to establish this correlation and its statistical significance. In a similar vein, the high frequency-of-occurrence of null subjects, which is observed in Set 3, predictably influences the system's precision, as the task of finding the antecedent of null subjects is more intricate.

	Set1	**Set2**	**Set3**
Precision	0,92	0,83	0,73
F	0,84	0,73	0,67
Mean number of sentences	7,78	9,77	29,73
Mean number of tokens	22,23	27,61	19,53
Mean number of mentions	0,40	0,18	0,13
Mean number of null subjects	0,10	0,12	0,24

Table 4 | System's precision in relation to the evaluation corpus size

5. Summary and future directions

In the present paper a coreference resolution system has been presented. To our knowledge, it is the first system that has been developed for the Greek language. It follows a domain-independent, rule-based approach that falls under the category of the entity-mention models. The system can handle all types of expressions, nominal and pronominal ones as well as null subjects.

High precision being mainly targeted, clustering of the coreferent expressions in a text is performed gradually through the employment of a set of morphological and syntactic criteria. The first system release, reported on here, has yielded promising results, having achieved precision of ~83%.

In parallel with the CR system, an annotation scheme has been developed, which apart from coreference, describes two types of relations between linguistic expressions: near identity and bridging. The scheme has been used for annotating a medium-size corpus of Greek texts mainly originating from the news domain.

Future directions involve handling of more demanding phenomena such as split antecedents or the detection of inter-sentential pronoun-antecedent chains. The anaphoricity check, through which expressions not referring to an entity are excluded from being possible candidates, could also be enhanced. Furthermore, exploiting semantic knowledge and/or the factor of context (cf. Recasens et al., 2013: 2) is expected to substantially improve the system's precision in detecting coreferential expressions. Finally, it is planned to experiment with different genres and domains in order to assess whether the domain/genre factor has an impact on performance, even though the system's rationale and implementation are domain-independent.

Acknowledgement

The research leading to the results above has been funded by the «POLYTROPON» project (ΟΠΣ: 448306, ΠΔΕ: 2013ΣΕ01380028).

References

Bagga, Amit, and Breck Baldwin. 1998. "Algorithms for Scoring Coreference Chains". In *Proceedings of the 1st International Conference on Language Resources and Evaluation (LREC'98), Workshop on Linguistic Coreference*, Granada, Spain, 563–566.

Bryl, Volha, Claudio Giuliano, Luciano Serafini, and Kateryna Tymoshenko. 2010. "Supporting Natural Language Processing with Background Knowledge: Coreference Resolution Case". In *Proceedings of the 9th International Semantic Web Conference (ISWC 2010)*, Shanghai, China, 80–95.

Durrett, Greg, and Dan Klein. 2013. "Easy Victories and Uphill Battles in Coreference Resolution". In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP 2013)*, Seattle, USA, 1971–1982.

Haghighi, Aria, and Dan Klein. 2009. "Simple Coreference Resolution with Rich Syntactic and Semantic Features". In *Proceedings of the EMNLP 2009, Conference on Empirical Methods in Natural Language Processing*, Singapore, 1152–1161.

Haghighi, Aria, and Dan Klein. 2010. "Coreference Resolution in a Modular, Entity-Centered Model". In *Proceedings of Human Language Technologies: The 11th Annual Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL 2010)*, Los Angeles, CA., USA, 385–393.

Iida, Ryu, Inui Kentaro, Takamura Hiroya, and Yuji Matsumoto. 2003. "Incorporating Contextual Cues in Trainable Models for Coreference Resolution". In *Proceedings of "The Computational Treatment of Anaphora" workshop held in conjunction with the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2003)*, Budapest, Hungary, 23–30.

Lappin, Shalom, and Herbert J. Leas. 1994. "An Algorithm for Pronominal Anaphora Resolution". *Computational Linguistics* 20(4): 535–561.

Lee, Heeyoung, Angel Chang, Yves Peirsman, Nathanael Chambers, Mihai Surdeanu, and Dan Jurafsky. 2013. "Deterministic Coreference Resolution Based on Entity-Centric, Precision-Ranked Rules". *Computational Linguistics* 39(4): 885–916.

Luo, Xiaoqiang, Abe Ittycheriah, Hongyan Jing, Nanda Kambhatla, and Salim Roukos. 2004. "A Mention-Synchronous Coreference Resolution Algorithm Based on the Bell Tree". In *Proceedings of the 42nd*

Annual Meeting of the Association for Computational Linguistics, Barcelona, Spain, 135–142.

Luo, Xiaoqiang. 2005. "On Coreference Resolution Performance Metrics". In *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing (HLT '05)*, Vancouver, British Columbia, Canada, 25–32.

Luo, Xiaoqiang, Sameer Pradhan, Marta Recasens, and Eduard Hovy. 2014. "An Extension of BLANC to System Mentions". In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Baltimore, Maryland, USA, 24–29.

Ng, Vincent, and Claire Cardie. 2002. "Improving Machine Learning Approaches to Coreference Resolution". In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL-02)*, Philadelphia, USA, 104–111.

Ng, Vincent. 2010. "Supervised Noun Phrase Coreference Research: The First Fifteen Years". In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL '10)*, Uppsala, Sweden, 1396–1411.

O'Connor, Brendan, and Michael Heilman. 2013. "ARKref: A Rule-Based Coreference Resolution System". ArXiv: 1310.1975. https://arxiv.org/abs/1310.1975

Ponzetto, Simone Paolo, and Michael Strube. 2007. "Knowledge Derived From Wikipedia For Computing Semantic Relatedness". *Journal of Artificial Intelligence Research* 30:181–212.

Recasens, Marta, and Antònia M. Martí. 2010. "AnCora-CO: Coreferentially Annotated Corpora for Spanish and Catalan". *Language Resources and Evaluation* 44(4):315–345.

Recasens, Marta, and Eduard Hovy. 2011. "BLANC: Implementing the Rand Index for Coreference Evaluation". *Natural Language Engineering* 17(4):485–510.

Recasens, Marta, Matthew Can, and Dan Jurafsky. 2013. "Same Referent, Different Words: Unsupervised Mining of Opaque Coreferent Mentions". In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2013)*, Atlanta, Georgia, USA, 897–906.

Soon, Wee Meng, Hwee Tou Ng, and Daniel Chung Yong Lim. 2001. "A Machine Learning Approach to Coreference Resolution of Noun Phrases". *Computational Linguistics* 27(4):521–544.

Uryupina, Olga, 2006. "Coreference Resolution with and without Linguistic Know-

ledge". In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, Genoa, Italy, 893–898.

Versley, Yannick, Simone Ponzetto, Massimo Poesio, Vladimir Eidelman, Alan Jern, Jason Smith, Xiaofeng Yang, and Alessandro Moschitti. 2008a. "BART: A Modular Toolkit for Coreference Resolution." In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, Marrakech, Morocco, 962–965.

Versley, Yannick, Alessandro Moschitti, Massimo Poesio, and Xiaofeng Yang. 2008b. "Coreference Systems based on Kernel Methods". In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, Manchester, UK, 961–968.

Vilain, Marc, John Burger, John Aberdeen, Dennis Connolly, and Lynette Hirschman. 1995. "A Model Theoretic Coreference Scoring Scheme". In *Proceedings of the 6th Conference on Message Understanding (MUC6 '95)*, Columbia, Maryland, 45–52.

Yang, Xiaofeng, Su Jian, Jun Lang, Tan Chew Lim, Ting Liu, and Sheng Li. 2008a. "An Entity-Mention Model for Coreference Resolution with Inductive Logic Programming. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT 2008)*, Columbus, Ohio, USA, 843–851.

Yang, Xiaofeng, Su Jian, GuoDong Zhou, and Tan Chew Lim. 2004. "An NP-Cluster Based Approach to Coreference Resolution". In *Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004)*, Geneva, Switzerland, 226–232.

Yang, Xiaofeng, Su Jian, and Tan Chew Lim. 2008b. "A Twin-Candidate Model for Learning-Based Anaphora Resolution". *Computational Linguistics* 34(3):327–356.