

Different German and English Coreference Resolution Models for Multi-domain Content Curation Scenarios

Ankit Srivastava, Sabine Weber, Peter Bourgonje^(✉), and Georg Rehm^(✉)

Language Technology Lab, DFKI GmbH, Alt-Moabit 91c, 10559 Berlin, Germany
{peter.bourgonje,georg.rehm}@dfki.de

Abstract. Coreference Resolution is the process of identifying all words and phrases in a text that refer to the same entity. It has proven to be a useful intermediary step for a number of natural language processing applications. In this paper, we describe three implementations for performing coreference resolution: rule-based, statistical, and projection-based (from English to German). After a comparative evaluation on benchmark datasets, we conclude with an application of these systems on German and English texts from different scenarios in digital curation such as an archive of personal letters, excerpts from a museum exhibition, and regional news articles.

1 Introduction to Coreference Resolution

Coreference resolution, the task of determining the mentions in a text, dialogue or utterance that refer to the same discourse entity, has been at the core of Natural Language Understanding since the 1960s. Owing in large part to publicly available annotated corpora, such as the Message Understanding Conferences (MUC) (Grishman and Sundheim 1996), Automatic Content Extraction (ACE) (Dodding et al. 2004), and OntoNotes¹, significant progress has been made in the development of corpus-based approaches to coreference resolution. Using coreference information has been shown to be useful in tasks such as question answering (Hartrumpf et al. 2008), summarisation (Bergler et al. 2003), machine translation (Miculicich Werlen and Popescu-Belis 2017), and information extraction (Zelenko et al. 2004).

Figure 1 shows a text consisting of three sentences and demonstrates the occurrence of two nouns and the mentions referring to them; *Prof. Hayes, Hayes, he* (shaded in yellow) and *I, me, Eric* (shaded in blue). The purpose of a coreference resolution system is to identify such chains of words and phrases referring to the same entity, often starting with (proper) noun phrases and referring pronouns.

The curation of digital information, has, in recent years, emerged as a fundamental area of activity for the group of professionals often referred to as knowledge workers. These knowledge workers are given the task to conduct research in

¹ <https://catalog.ldc.upenn.edu/LDC2013T19>.

a particular domain in a very limited time frame. The output of their work is used by newspaper agencies to create articles; museums to construct new exhibitions on a specific topic; TV stations to generate news items. Owing to the diversity of tasks and domains they have to work in, knowledge workers face the challenge to explore potentially large multimedia document collections and quickly grasp key concepts and important events in the domain they are working in. In an effort to help them, we can automate some processes in digital curation such as the identification of named entities and events. This is the primary use case for our paper as coreference resolution plays a significant role in disambiguation as well as harnessing a larger number of entities and events. For example, as seen in Fig. 1, after linking *He* and *Hayes* with *Prof. Hayes*, the knowledge worker gets more information to work with.

#1: Wednesday night: Dinner at Prof. Hayes and last lecture, which I had worked over to a proper climax.
 #2: Hayes did said: the best series of lectures he ever heard!
 #3: He calls me Eric now.

Fig. 1. Example of coreference occurrence in English text. Source: Mendelsohn letters dataset (Bienert and de Wit 2014) (Color figure online)

While many coreference systems exist for English (Raghunathan et al. 2010; Kummerfeld and Klein 2013; Clark and Manning 2015, 2016), a freely available² competitive tool for German is still missing. In this paper, we describe our forays into developing a German coreference resolution system. We attempt to adapt the Stanford CoreNLP (Manning et al. 2014) Deterministic (rule-based) Coreference Resolution approach (Raghunathan et al. 2010; Lee et al. 2013) as well as the Stanford CoreNLP Mention Ranking (statistical) model (Clark and Manning 2015) to German. We also experiment with projection-based implementation, i.e., using Machine Translation and English coreference models to achieve German coreference resolution.

The main goals of this paper are:

- To evaluate pre-existing English and German coreference resolution systems
- To investigate the effectiveness of performing coreference resolution on a variety of out-of-domain texts in both English and German (outlined in Sect. 4) from digital curation scenarios.

After a brief overview of previous approaches to coreference resolution in English and German (Sect. 2), we describe implementations of three approaches

² Coreference resolution is language resource dependent and therefore by “freely available” we imply a toolkit which in its entirety (models, dependencies) is available for commercial as well as research purposes.

to German coreference resolution (Sect. 3): the deterministic sieve-based approach, a machine learning-based system, and a English-German crosslingual projection-based system. This is followed by a discussion on applications of coreference (Sect. 4) and concluding notes on the current state of our coreference resolution systems for digital curation scenarios (Sect. 5).

2 Summary of Approaches to Coreference Resolution

A number of paradigms (rule-based, knowledge-rich, supervised and unsupervised learning) have been applied in the design of coreference resolution systems for several languages with regard to whole documents, i.e., to link all mentions or references of an entity within an entire document. While there have been several works giving a comprehensive overview of such approaches (Zheng et al. 2011; Stede 2011), we focus on coreference resolution for German and English and summarise some of the systems.

There have been several attempts at performing coreference resolution for German documents and building associated systems.³ *CorZu* (Tugener 2016) is an incremental entity-mention system for German, which addresses issues such as underspecification of mentions prevalent in certain German pronouns. While it is freely available under the GNU General Public License, it depends on external software and their respective data formats such as a dependency parser, tagger, and morphological analyser, making it difficult to reimplement it.

BART, the Beautiful/Baltimore Anaphora Resolution Toolkit (Versley et al. 2008), is a modular toolkit for coreference resolution which brings together several preprocessing and syntactic features and maps it to a machine learning problem. While it is available for download as well as a Web Service, there are external dependencies such as the Charniak Reranking Parser. Definite noun matching is resolved via head string matching, achieving an F-score of 73% (Versley 2010). It has been successfully tested on the German TüBa-D/Z treebank (Telljohann et al. 2004) with a claimed F-score of 80.2% (Broscheit et al. 2010).

Definite noun matching cannot be solved via string matching in the domain of newspaper articles. Approximately 50% of the definite coreferent Noun Phrases (NPs) can be resolved using head string matching (Versley 2010). Versley (2010) also used hypernym look-up and various other features to achieve an F-score of 73% for definite anaphoric NPs. Broscheit et al. (2010) claim to get an F1 score of 80.2 on version 4 of the TüBa D/Z coreference corpus using BART.

The goal of the SemEval 2010 Shared Task 1 (Recasens et al. 2010) was to evaluate and compare automatic coreference resolution systems for six different languages, among them German, in four evaluation settings and using four different metrics. The training set contained 331,614 different tokens taken from the TüBa-D/Z data set (Telljohann et al. 2004). Only two of the four competing systems achieved F-scores over 40%, one of them being the BART system

³ For an overview of the development of German coreference systems, see Tugener (2016).

mentioned above. We use the same dataset and evaluation data to train our statistical system in Sect. 3.1.

Departing from the norm of building mention pairs, one system implemented a mention-entity approach and produced an F-score of 61.49% (Klenner et al. 2010).

The *HotCoref* system for German (Roesiger and Riester 2015) focused on the role of prosody for coreference resolution and used the DIRNDL corpus (Björkelund et al. 2014) for evaluation, achieving F-scores of 53.63% on TüBa-D/Z (version 9) and 60.35% on the SemEval shared task data.

Another system (Krug et al. 2015) adapted the Stanford sieve approach (Lee et al. 2013) for coreference resolution in the domain of German historic novels and evaluated it against a hand annotated corpus of 48 novel fragments with approximately 19,000 character references in total. An F1 score of 85.5 was achieved. We also adapt the Stanford Sieve approach in Sect. 3, with the aim of developing an open-domain German coreference resolution system.

In case of the English coreference resolution, we employ the Stanford CoreNLP implementations. There is a large body of work for coreference resolution in English. While the sieve-based approach (Raghunathan et al. 2010) is a prime example of rule-based coreference resolution, other approaches such as the Mention-Rank model (Clark and Manning 2015) and Neural model (Clark and Manning 2016) have been shown to outperform it.

3 Three Implementations

In this section, we describe the three models of coreference resolution.

- Rule-based (Multi-Sieve Approach): English, German
- Statistical (Mention Ranking Model): English, German
- Projection-based (Crosslingual): coreference for German using English models.

3.1 Rule-Based Approach

For the English version, we employ the deterministic multi-pass sieve-based (open-source) Stanford CoreNLP system (Manning et al. 2014). For the German version, we develop an in-house system and name it **CoRefGer-rule**⁴.

The Stanford Sieve approach is based on the idea of an annotation pipeline with coreference resolution being one of the last steps. The processing steps include sentence splitting, tokenisation, constituency and dependency parsing, and extraction of morphological data. In our system **CoRefGer-rule**, we also perform Named Entity Recognition.

What is typical for the Stanford sieve approach is starting with all noun phrases and pronominal phrases in the whole document and then deciding how

⁴ Source code available at <https://github.com/dkt-projekt/e-NLP/tree/master/src/main/java/de/dkt/eservices/ecorenlp/modules>.

to cluster them together, so that all the noun phrases referring to the same extratextual entity are in the same coreference chain. The sieves can be described as a succession of independent coreference models. Each of them selects candidate mentions and puts them together. The number of these sieves can be different depending on the task. Seven sieves are proposed for an English coreference system (Raghunathan et al. 2010), while eleven sieves are implemented for the task of finding coreference in German historic novels (Krug et al. 2015). We have currently implemented six of the seven sieves from the English system and will include additional ones in future versions of the system.

Sieve 1: Exact Match. With an exact match, noun phrases are extracted from the parse tree. Then, in a sliding window of five sentences, all noun phrases in this window are compared to each other. If they match exactly this leads to the creation of a new coreference chain. We use stemming so that minimally different word endings and differences in the article are taken into account (Table 1).

Table 1. Example for exact match

Text (de):	Barack Obama besuchte Berlin Am Abend traf Barack Obama die Kanzlerin
Coref:	[Barack Obama, Barack Obama]

We also account for variations in endings such as “des Landesverbandes” and “des Landesverbands der AWO” or “der Hund” and “des Hundes”, and between definite and indefinite articles such as “einen Labrador” and “der Labrador”.

Sieve 2: Precise Constructs. This is an implementation of precise constructs like appositive constructs, predicate nominative, role appositive and relative pronouns. Due to the different tree tags a direct application of Stanford NLP algorithms was not possible. Also missing acronym and demonym lists for German posed a challenge in completing this sieve. We therefore translated the corresponding English lists ⁵ into German and used them in our approach.

Sieves 3, 4, 5: Noun Phrase Head Matching. The noun phrase head matching we use is different from the one proposed in Raghunathan et al. (2010). They claim that naive matching of heads of noun phrases creates too many spurious links, because it ignores incompatible modifiers like “Yale University” and “Harvard University”. Those two noun phrases would be marked as coreferent, because the head of both is “University”, although the modifiers make it clear that they refer to different entities. This is why a number of other constraints are proposed. In order to utilise them we implement a coreference chain building mechanism. For example, there is a notion of succession of the words when

⁵ https://en.wikipedia.org/wiki/List_of_adjectival_and_demonymic_forms_of_place_names.

chaining them together, so we cannot match head nouns or noun phrases in the antecedent cluster.

We also employ stemming in noun phrases, so we match entities such as “AWO-Landesverbands” with “Landesverband”, “Geschäftsführer” and “Geschäftsführers.”

We also implement the sieves called “Variants of head matching” and “relaxed head matching” which require sophisticated coreference chaining.

Sieve 6: Integration of Named Entity Recognition. We use an in-house Named Entity Recognition engine based on DBpedia-Spotlight⁶, that is also applied in the current version of the coreference resolution system (for example, to deal with the above mentioned “Yale University” vs. “Harvard University” issue).

German Specific Processing. Our implemented sieves include naive stemming, which means that words that vary in a few letters at the end are still considered as matching due to different case markers in German. The same holds for definite and indefinite articles, which are specific for German. Noun phrases are considered as matching although they have different articles.

An important component that is not implemented but plays a big role in coreference resolution is morphological processing for acquiring gender and number information. This component would make it possible to do pronoun matching other than our current method of merely matching the pronouns that are the same.

3.2 Statistical Approach

While we developed the rule-based approach (CoRefGer-rule), we also adapted the Stanford CoreNLP statistical system based on the Mention Ranking model (Clark and Manning 2015). We trained our coreference system on the TüBa-D/Z (Telljohann et al. 2004), and evaluated on the same dataset as SemEval 2010 Task 1 (Recasens et al. 2010). We named the system **CoRefGer-stat**. The system uses a number of features described below:

- Distance features: the distance between the two mentions in a sentence, number of mentions
- Syntactic features: number of embedded NPs under a mention, Part-Of-Speech tags of the first, last, and head word (based on the German parsing models included in the Stanford CoreNLP, Rafferty and Manning 2008)
- Semantic features: named entity type, speaker identification
- Lexical Features: the first, last, and head word of the current mention.

While the machine learning approach enables robustness and saves time in constructing sieves, the application is limited to the news domain, i.e., the domain of the training data.

⁶ <https://github.com/dbpedia-spotlight/dbpedia-spotlight>.

3.3 Projection-Based Approach

In this section we outline the projection-based approach to coreference resolution. This approach is usually implemented in a low-resource language scenario, i.e., if sufficient language resources and training data are not available. Developing a coreference resolution system for any new language is cumbersome due to the variability of coreference phenomena in different languages as well as availability of high-quality language technologies (mention extraction, syntactic parsing, named entity recognition).

Crosslingual projection is a mechanism that allows transferring of existing methods and resources from one language (e.g., English) to another language (e.g., German). While crosslingual projection has demonstrated considerable success in various NLP applications like POS tagging and syntactic parsing, it has been less successful in coreference resolution, performing with 30% less precision than monolingual variants (Grishina and Stede 2017).

The projection-based approach can be implemented in one of the following two ways:

- Transferring models: Computing coreference on text in English, and projecting these annotations on parallel German text via word alignments in order to obtain German coreference model
- Transferring data: Translating German text to English, computing coreference on translated English text using English coreference model and then projecting the annotations back on to the original text via word alignment.

The “Transferring data” approach involved less overhead because new language coreference models are not generated, and proved to be more effective. We have therefore used this approach in our experiments and name it **CoRefGerproj** system.

4 Evaluation and Case Studies

We are interested in applying reliable and robust coreference resolution for both English and German on a variety of domains from digital curation scenarios such as digital archives, newspaper reports, museum exhibits (Bourgonje et al. 2016, Rehm and Sasaki 2016):

- Mendelsohn Letters Dataset (German and English): The collection (Bienert and de Wit 2014) contains 2,796 letters, written between 1910 and 1953, with a total of 1,002,742 words on more than 11,000 sheets of paper; 1,410 of the letters were written by Erich and 1,328 by Luise Mendelsohn. Most are in German (2,481), the rest is written in English (312) and French (3).
- Research excerpts for a museum exhibition (English): This is a document collection retrieved from online archives: Wikipedia, archive.org, and Project Gutenberg. It contains documents related to Vikings; the content of this collection has been used to plan and to conceptualise a museum in Denmark.

Table 2. Summary of curations datasets.

Corpora	Language	Documents	Words	Domain
Mendelsohn	DE	2,501	699,213	Personal letters
Mendelsohn	EN	295	21,226	Personal letters
Vikings	EN	12	298,577	Wikipedia and E-books
News	DE	1,037	716,885	News articles and summaries

- Regional news stories (German): This consists of a general domain regional news collection in German. It contains 1,037 news articles, written between 2013 and 2015.

The statistics for these corpora and the standard benchmark sets are summarised in Table 2. Robustness can only be achieved if we limit the scope and coverage of the approach, i.e., if we keep the coreference resolution systems simple and actually implementable. In our use cases, a few correctly identified mentions are better than hundreds of false positives.

Evaluation is done on several datasets (standard datasets for benchmarking): CoNLL 2012 for English and SemEval 2010 for German. Our goal is to determine the optimal coreference system for coreference resolution on English and German texts from digital curation scenarios (out-of-domain).

Table 3 shows the results of evaluation on CoNLL 2012 (Pradhan et al. 2012) English dataset. Two evaluation measures are employed: MUC (Vilain et al. 1995) and B-cubed (Bagga and Baldwin 1998). The MUC metric compares links between mentions in the key chains to links in the response chains. The B-cubed metric evaluates each mention in the key by mapping it to one of the mentions in the response and then measuring the amount of overlapping mentions in the corresponding key and response chains.

Table 3. Summary of evaluation of 2 coreference resolution systems on English CoNLL 2012 task across two F-1 evaluation measures.

System	MUC	B-cube
BART	45.3	64.5
Sieve	49.2	45.3
Statistical	56.3	50.4
Neural	60.0	56.8

Note that these accuracies are lower than those reported in CoNLL shared task, because the systems employed are dependent on taggers, parsers and named entity recognizers of Stanford CoreNLP and not gold standard as employed in the shared task. While we do not have any gold standard for our digital curation use-cases, a manual evaluation of a small subset of documents shows sieve-based approach to perform slightly better than the state-of-the-art statistical and neural models, most likely owing to out-of-domain applications.

For German, we experimented with different settings of the in-house CoRefGer-rule system. In Table 4, we demonstrate the performance of our Multi-Sieve (rule-based) approach on a 5000-word subset of the TüBa-D/Z corpus (Telljohann et al. 2004) using different settings of modules as follows:

- Setting 1: Whole System with all 6 sieves in place
- Setting 2: Contains all mentions but no coreference links
- Setting 3: Setting 1 minus the module that is deleting any cluster that does not contain a single mention that has been recognized as an entity
- Setting 4: Setting 1 with the module that is deleting any cluster that does not contain a single mention that has been recognized as an entity executed after the sieves have been applied.

Setting 2 assumes that our system obtained all the correct mentions and therefore tests the effectiveness of the coreference linking module only. However this setting will not work in real-life scenarios like the digital curation use cases unless we have hand-annotated corpora.

Table 4. Module-based evaluation of German sieve-based coreference (CoRefGer-rule) on different configurations across two F-1 evaluation measures.

System	MUC	B-cube
Setting 1	54.4	11.2
Setting 2	70.5	23.1
Setting 3	58.9	15.0
Setting 4	56.1	12.0

In Table 5, we present German coreference resolution results on the test set of SemEval 2010 Task 1. We also compare the performance of our three systems (CoRefGer-rule, CoRefGer-stat, CoRefGer-proj) to one other system: CorZU (Tuggener 2016). CoRefGer-stat, CoRefGer-rule, and CoRefGer-proj are the three systems we developed in this paper. Since the sieve-based approach lacks a morphological component currently, it underperforms. An error analysis of the statistical and projection-based system reveals that several features were not sufficiently discriminating for German models. We believe completing the remaining sieve will help us in training better syntactic and semantic features for the statistical system as well.

While there is no gold standard for any of our datasets from digital curation use cases, we nevertheless applied our English and German coreference resolution systems, as shown in Table 6. The sieve-based systems tend to give the best results (shown in the Table) while the statistical, neural and projection-based yield nearly 10% less entity mentions. We leave for future work a deeper investigation into this though we believe that interfacing with lexical resources such as those from WordNet may help ameliorate the out-of-domain issues.

Table 5. Summary of evaluation of coreference resolution systems on German SemEval 2010 task across 2 F-1 evaluation measures.

System	MUC	B-cube
CorZu	60.1	58.9
CoRefGer-rule	50.2	63.3
CoRefGer-stat	40.1	45.3
CoRefGer-proj	35.9	40.3

Table 6. Summary of the percentage of mentions (based on total number of words) on curation datasets for which we do not have a gold standard

Dataset	Sents.	Words	Mentions
Mendelsohn EN	21K	109K	48%
Mendelsohn DE	34K	681K	26%
Vikings EN	39K	310K	49%
News Stories DE	53K	369K	25%

4.1 Add-On Value of Coreference Resolution to Digital Curation Scenarios

Consider the following sentence:

“Then came Ray Brock for dinner. On **him** I will elaborate after my return or as soon as a solution pops up on my “Klappenschrank”. Naturally, **he** sends **his** love to Esther and **his** respects to you.”

A model or dictionary can only spot “Ray Brock”, but, “him”, “he” and “his” also refer to this entity. With the aid of coreference resolution, we can increase the recall for named entity recognition as well as potentially expand the range for event detection.

The algorithm for coreference-enabled NLP technologies is as follows:

- Input a text document, and run coreference resolution on it
- With the aid of the above, replace all occurrences of pronouns with the actual noun in full form, such that “he” and “his” are replaced with “Ray Brock” and “Ray Brock’s” respectively
- Run a NLP process such as Named Entity Recognition on the new document and compare with a run without the coreference annotations.

A preliminary computation of the above algorithm shows a marked improvement on the number of entities identified (27% more coverage) by an off-the-shelf Named Entity Recogniser.

5 Conclusions and Future Work

We have performed coreference resolution in both English and German on a variety of text types and described several competing approaches (rule-based, statistical, knowledge-based).

Number and gender information is one of the core features that any coreference system uses. A major deficiency for our German rule-based system described in Sect. 3.1 is the lack of interfacing with a morphological analyser, which we leave for future work.

An interesting sieve that can also be adapted from the paper about German coreference resolution in historic novels is the semantic pass where synonyms taken from GermaNet are also taken into account for matching. They handle speaker resolution and pronoun resolution in direct speech, which makes up their tenth and eleventh sieve. They do so by using handcrafted lexico-syntactic patterns. If these patterns are only specific for their domain or if they can also be successfully applied to other domains is a point for further research.

Overall, we were able to annotate our multi-domain datasets with coreference resolution. We will be investigating how much these annotations help knowledge workers in their curation use cases.

In conclusion, we have determined that the deterministic rule-based systems, although not state-of-the-art are better choices for our out-of-domain use cases.

Acknowledgments. We would like to thank the anonymous reviewers for their insightful and helpful comments. The project Digitale Kuratierungstechnologien (DKT) is supported by the German Federal Ministry of Education and Research (BMBF), Unternehmen Region, instrument Wachstumskern-Potenzial (No. 03WKP45). More details: <http://www.digitale-kuratierung.de>.

References

- Bagga, A., Baldwin, B.: Algorithms for scoring coreference chains. In: The First International Conference on Language Resources and Evaluation Workshop on Linguistics Coreference, pp. 563–566 (1998)
- Bergler, S., Witte, R., Khalife, M., Li, Z., Rudzicz, F.: Using knowledge-poor coreference resolution for text summarization. In: Proceedings of the Document Understanding Conference (DUC 2003) (2003). <http://duc.nist.gov/pubs/2003final.papers/concordia.final.pdf>
- Bienert, A., de Wit, W. (eds.): EMA - Erich Mendelsohn Archiv. Der Briefwechsel von Erich und Luise Mendelsohn 1910–1953. Kunstbibliothek - Staatliche Museen zu Berlin and The Getty Research Institute, Los Angeles, March 2014. <http://ema.smb.museum>. With contributions from Regina Stephan and Moritz Wullen, Version March 2014
- Björkelund, A., Eckart, K., Riester, A., Schaufler, N., Schweitzer, K.: The extended dirndl corpus as a resource for coreference and bridging resolution. In: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014). European Language Resources Association (ELRA) (2014). <http://www.lrec-conf.org/proceedings/lrec2014/pdf/891.Paper.pdf>

- Bourgonje, P., Moreno-Schneider, J., Nehring, J., Rehm, G., Sasaki, F., Srivastava, A.: Towards a platform for curation technologies: enriching text collections with a semantic-web layer. In: Sack, H., Rizzo, G., Steinmetz, N., Mladeni, D., Auer, S., Lange, C. (eds.) *The Semantic Web: ESWC 2016 Satellite Events*. LNCS, pp. 65–68. Springer International Publishing, Heidelberg (2016). https://doi.org/10.1007/978-3-319-47602-5_14. ISBN 978-3-319-47602-5
- Broscheit, S., Ponzetto, S.P., Versley, Y., Poesio, M.: Extending bart to provide a coreference resolution system for German. In: *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC 2010)*. European Languages Resources Association (ELRA) (2010). <http://aclweb.org/anthology/L10-1347>
- Clark, K., Manning, C.D.: Entity-centric coreference resolution with model stacking. In: *Association for Computational Linguistics (ACL)* (2015)
- Clark, K., Manning, C.D.: Deep reinforcement learning for mention-ranking coreference models. In: *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 2256–2262. Association for Computational Linguistics (2016). <http://aclweb.org/anthology/D16-1245>
- Doddington, G., Mitchell, A., Przybocki, M., Ramshaw, L., Strassel, S., Weischedel, R.: The automatic content extraction (ACE) program tasks, data, and evaluation. In: *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004)*. European Language Resources Association (ELRA) (2004). <http://aclweb.org/anthology/L04-1011>
- Grishina, Y., Stede, M.: Multi-source annotation projection of coreference chains: assessing strategies and testing opportunities. In: *Proceedings of the 2nd Workshop on Coreference Resolution Beyond OntoNotes (CORBON 2017)*, Valencia, Spain, pp. 41–50. Association for Computational Linguistics, April 2017. <http://www.aclweb.org/anthology/W17-1506>
- Grishman, R., Sundheim, B.: Message understanding conference- 6: a brief history. In: *COLING 1996 Volume 1: The 16th International Conference on Computational Linguistics (1996)*. <http://aclweb.org/anthology/C96-1079>
- Hartrumpf, S., Glöckner, I., Leveling, J.: Coreference resolution for questions and answer merging by validation. In: Peters, C., Jijkoun, V., Mandl, T., Müller, H., Oard, D.W., Peñas, A., Petras, V., Santos, D. (eds.) *CLEF 2007*. LNCS, vol. 5152, pp. 269–272. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-85760-0_32
- Klenner, M., Tuggener, D., Fahrni, A., Sennrich, R.: Anaphora resolution with real preprocessing. In: Loftsson, H., Rögnavaldsson, E., Helgadóttir, S. (eds.) *NLP 2010*. LNCS (LNAI), vol. 6233, pp. 215–225. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-14770-8_25
- Krug, M., Puppe, F., Jannidis, F., Macharowsky, L., Reger, I., Weimar, L.: Rule-based coreference resolution in German historic novels. In: *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, Denver, Colorado, USA, pp. 98–104. Association for Computational Linguistics, June 2015. <http://www.aclweb.org/anthology/W15-0711>
- Kummerfeld, J.K., Klein, D.: Error-driven analysis of challenges in coreference resolution. In: *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, Seattle, Washington, USA, pp. 265–277, October 2013. <http://www.aclweb.org/anthology/D13-1027>

- Lee, H., Chang, A., Peirsman, Y., Chambers, N., Surdeanu, M., Jurafsky, D.: Deterministic coreference resolution based on entity-centric, precision-ranked rules. *Computational Linguistics* **39**(4) (2013). https://doi.org/10.1162/COLI_a.00152, <http://aclweb.org/anthology/J13-4004>
- Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S.J., McClosky, D.: The Stanford CoreNLP natural language processing toolkit. In: *Association for Computational Linguistics (ACL) System Demonstrations*, pp. 55–60 (2014). <http://www.aclweb.org/anthology/P/P14/P14-5010>
- Werlen, L.M., Popescu-Belis, A.: Proceedings of the 2nd Workshop on Coreference Resolution Beyond OntoNotes (CORBON 2017), Chapter Using Coreference Links to Improve Spanish-to-English Machine Translation, pp. 30–40. *Association for Computational Linguistics* (2017). <http://aclweb.org/anthology/W17-1505>
- Pradhan, S., Moschitti, A., Xue, N., Uryupina, O., Zhang, Y.: Conll-2012 shared task: modeling multilingual unrestricted coreference in ontototes. In: *Joint Conference on EMNLP and CoNLL - Shared Task, CoNLL 2012*, Stroudsburg, PA, USA, pp. 1–40. *Association for Computational Linguistics* (2012). <http://dl.acm.org/citation.cfm?id=2391181.2391183>
- Rafferty, A.N., Manning, C.D.: Parsing three German treebanks: lexicalized and unlexicalized baselines. In: *Proceedings of the Workshop on Parsing German, PaGe 2008*, Stroudsburg, PA, USA, pp. 40–46. *Association for Computational Linguistics* (2008). ISBN 978-1-932432-15-2. <http://dl.acm.org/citation.cfm?id=1621401.1621407>
- Raghunathan, K., Lee, H., Rangarajan, S., Chambers, N., Surdeanu, M., Jurafsky, D., Manning, C.: A multi-pass sieve for coreference resolution. In: *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pp. 492–501. *Association for Computational Linguistics* (2010). <http://aclweb.org/anthology/D10-1048>
- Recasens, M., Màrquez, L., Sapena, E., Martí, A.M., Taulé, M., Hoste, V., Poesio, M., Versley, Y.: Semeval-2010 task 1: coreference resolution in multiple languages. In: *Proceedings of the 5th International Workshop on Semantic Evaluation*, pp. 1–8. *Association for Computational Linguistics* (2010). <http://aclweb.org/anthology/S10-1001>
- Rehm, G., Sasaki, F.: Digital curation technologies. In: *Proceedings of the 19th Annual Conference of the European Association for Machine Translation (EAMT 2016)*, Riga, Latvia, May 2016, In print
- Roesiger, I., Riester, A.: Using prosodic annotations to improve coreference resolution of spoken text. In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pp. 83–88. *Association for Computational Linguistics* (2015). <https://doi.org/10.3115/v1/P15-2014>, <http://aclweb.org/anthology/P15-2014>
- Stede, M.: *Discourse Processing*. *Synthesis Lectures in Human Language Technology*, vol. 15. Morgan and Claypool, San Rafael (2011)
- Telljohann, H., Hinrichs, E., Kübler, S.: The tüba-d/z treebank: annotating German with a context-free backbone. In: *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004)*. *European Language Resources Association (ELRA)* (2004). <http://aclweb.org/anthology/L04-1096>
- Tuggener, D.: *Incremental coreference resolution for German*. Ph.D. thesis, University of Zurich (2016)
- Versley, Y.: *Resolving coreferent bridging in German newspaper text*. Ph.D. thesis, Universität Tübingen (2010)

- Versley, Y., Ponzetto, P.S., Poesio, M., Eidelman, V., Jern, A., Smith, J., Yang, X., Moschitti, A.: Bart: a modular toolkit for coreference resolution. In: Proceedings of the ACL-08: HLT Demo Session, pp. 9–12. Association for Computational Linguistics (2008). <http://aclweb.org/anthology/P08-4003>
- Vilain, M., Burger, J., Aberdeen, J., Connolly, D., Hirschman, L.: A model-theoretic coreference scoring scheme. In: Proceedings of the 6th Conference on Message Understanding, MUC6 1995, Stroudsburg, PA, USA, pp. 45–52 (1995). Association for Computational Linguistics. ISBN 1-55860-402-2. <https://doi.org/10.3115/1072399.1072405>
- Zelenko, D., Aone, C., Tibbetts, J.: Proceedings of the Conference on Reference Resolution and Its Applications, Chapter Coreference Resolution for Information Extraction (2004). <http://aclweb.org/anthology/W04-0704>
- Zheng, J., Chapman, W., Crowley, R., Guergana, S.: Coreference resolution: a review of general methodologies and applications in the clinical domain. *Journal of Biomedical Informatics* 44(6) (2011). <https://doi.org/10.1016/j.jbi.2011.08.006>, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3226856/>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

