Events are Not Simple: Identity, Non-Identity, and Quasi-Identity

Eduard Hovy
Language Technology Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213, USA
hovy@cmu.edu

Teruko Mitamura
Language Technology Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213, USA
teruko@cs.cmu.edu

Felisa Verdejo
E.T.S.I. Informática, UNED
C/ Juan del Rosal, 16
(Ciudad Universitaria)
28040 Madrid, Spain
felisa@lsi.uned.es

Jun Araki
Language Technology Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213, USA
junaraki@cs.cmu.edu

Andrew Philpot
Information Sciences Institute
University of Southern California
4676 Admiralty Way
Marina del Rey, CA 90292, USA
philpot@isi.edu

Abstract[1]

Despite considerable theoretical and computational work on coreference, deciding when two entities or events are identical is very difficult. In a project to build corpora containing coreference links between events, we have identified three levels of event identity (full, partial, and none). Event coreference annotation on two corpora was performed to validate the findings.

1 The Problem of Identity

Last year we had HLT in Montreal, and this year we did it in Atlanta.

Does the "did it" refer to the same conference or a different one? The two conferences are not identical, of course, but they are also not totally unrelated—else the "did it" would not be interpretable.

When creating text, we treat instances of entities and events as if they are fixed, well-described, and well-understood. When we say "that boat over there" or "Mary's wedding next month", we assume the reader creates a mental representation of the referent, and we proceed to refer to it without further thought.

However, as has been often noted in theoretical studies of semantics, this assumption is very problematic (Mill, 1872; Frege 1892; Guarino, 1999). Entities and (even more so) events are complex composite phenomena in the world, and they undergo change.

Since nobody has complete knowledge, the author's mental image of the entity or event in question might differ from the reader's, and from the truth. Specifically, the properties the author assumes for the event or entity might not be the ones the reader assumes. This difference has deep consequences for the treatment of the semantic meaning of a text. In particular, it fundamentally affects how one must perform coreference among entities or events.

As discussed in Section 6, events have been the focus of study in both Linguistics and NLP (Chen and Ji, 2009; Bejan and Harabagiu, 2008, 2010; Humphreys et al., 1997). Determining when two event mentions in text corefer is, however, an unsolved problem[2]. Past work in NLP has avoided some of the more complex problems by considering only certain types of coreference, or by simply ignoring the major problems. The results have been partial, or inconsistent, annotations.

In this paper we describe our approach to the problem of coreference among events. In order to build a corpus containing event coreference links that is annotated with high enough inter-annotator agreement to be useful for machine learning, it has proven necessary to create a model of event identity that is more elaborate than is usually assumed in the NLP literature, and to formulate quite specific definitions for its central concepts.

[1] This work was supported by grants from DARPA and NSF, as well as by funding that supported Prof. M. Felisa Vedejo from UNED Madrid.

[2] In this work, we mean both events and states when we say "event". A state refers to a fixed, or regularly changing, configuration of entities in the world, such as "it is hot" or "he is running". An event occurs when there is a change of state in the world, such as "he stops running" or "the plane took off".

Proceedings of the The 1st Workshop on EVENTS: Definition, Detection, Coreference, and Representation, pages 21–28,
Atlanta, Georgia, 14 June 2013.

Event coreference is the problem of determining when two mentions in a text refer to the 'same' event. Whether or not the event actually occurred in reality is a separate issue; a text can describe people flying around on dragons or broomsticks. While the events might be actual occurrences, hypothesized or desired ones, etc., they exist in the text as *Discourse Elements* (DEs), and this is what we consider in this work.

Each DE is referred to (explicitly or implicitly) in the text by a *mention*, for example "destroy", "the attack", "that event", or "it". But it is often unclear whether two mentions refer to the same DE or to closely related ones, or to something altogether different. The following example illustrates two principal problems of event coreference:

> While Turkish troops have been **fighting_E.1** a Kurdish faction in northern Iraq, two other Kurdish groups have been **battling_E.2** each other.
>
> A radio station **operated_E.3** by the Kurdistan Democratic Party **said_E.4** the party's forces **attacked_E.5** positions of the Patriotic Union of Kurdistan on Monday in the Kurdish region's capital Irbil.
>
> The Voice of Iraqi Kurdistan radio, **monitored_E.6** by the British Broadcasting Corp., **said_E.7** more than 80 Patriotic Union fighters **were killed_E.8** and at least 150 **wounded_E.9**.
>
> The **fighting_E.10** was also **reported_E.11** by a senior Patriotic Union official, Kusret Rasul Ali, who **said_E.12** PUK forces **repelled_E.13** a large KDP **attack_E.14**.
>
> …
>
> Ali **claimed_E.16** that 300 KDP fighters were **killed_E.17** or **wounded_E.18** and only 11 Patriotic Union members **died_E.19**.

Problem 1: Partial event overlap. Event E.2, "battling each other", refers to an ongoing series of skirmishes between two Kurdish groups, the KDP and the PUK. Since one of these battles, where the KDP attacked positions of the PUK, is E.5, it is natural to say that E.2 and E.5 corefer. However, E.2 clearly denotes other battles as well, and therefore E.5 and E.2 cannot fully corefer. In another example, event E.8 refers to the killing of a number of soldiers as part of this fight E.5, and event E.9 to the wounding of others. Both events E.8 and E.9 constitute an intrinsic part of the attack E.5, and hence corefer to it, but are each only part of E.5, and hence neither can fully corefer to it.

Problem 2: Inconsistent reporting. This news fragment contains two reports of the fight: E.5 and E.10. Since E.10 describes E.5 from the perspective of a senior PUK official, it should corefer to E.5. But where the KDP's report claims more than 80 PUK fighters killed (event E.8, part of E.5), the PUK official said that only 11 PUK members died (event E.19, part of E.10). Without taking into account the fact that the two killing events are reports made by different speakers, it would not be possible to recognize them as coreferent.

Examples of partial event overlap and inconsistent reporting are common in text, and occur as various types. In our work, we formally recognize partial event overlap, calling it *partial event identity*, which permits different degrees and types of event coreference. This approach simplifies the coreference problem and highlights various inter-event relationships that facilitates grouping events into 'families' that support further analysis and combination with other NLP system components.

In this paper, we introduce the idea that there are three degrees of event identity: fully identical, quasi-identical, and fully independent (not identical). Full identity reflects in full coreference and quasi-identity in partial coreference. Fully independent events are singletons.

Our claims in this paper are:

- Events, being complex phenomena, can corefer fully (identity) or partially (quasi-identity).
- Event coreference annotation is considerably clarified when partial coreference is allowed.
- A relatively small fixed set of types of quasi-identity suffices to describe most of them.
- Different domains and genres highlight different subsets of these quasi-identity types.
- Different auxiliary knowledge sources and texts are relevant for different types.

2 Types of Full and Partial Identity

Def: Two mentions **fully corefer** if their activity/event/state DE is identical in all respects, as far as one can tell from their occurrence in the text. (In particular, their agents, location, and time are identical or compatible.) One can distinguish several types of identity, as spelled out below.

Def: Two mentions **partially corefer** if activity/event/state DE is quasi-identical: most aspects are the same, but some additional information is provided for one or the other that is not shared. There are two principal types of quasi-identity, as defined below.

Otherwise, two mentions **do not corefer**.

2.1 Full Identity

Mention1 is identical to mention2 iff there is no semantic (meaning) difference between them. Just one DE, and exactly the same aspects of the DE, are understood from both mentions in their contexts. It is possible to replace the one mention with the other without any semantic change (though some small syntactic changes might be required to ensure grammaticality). Note that mention2 may contain less detail than mention1 and remain identical, if it carries over information from mention1 that is understood / inherited from the context. However, when mention2 provides more or new information not contained in mention1 or naturally inferred for it, then the two are no longer identical. Usually, exact identity is rare within a single text, but may occur more often across texts. We identify the following types:

1. **Lexical identity**: The two mentions use exactly the same senses of the same word(s), including derivational words (e.g., "destroy", "destruction").

2. **Synonym**: One mention's word is a synonym of the other's word.

3. **Wide-reading**: One mention is a synonym of the wide reading of the other (defined below, under Quasi-identity:Scriptal). For example, in "the attack(E1) took place yesterday. The bombing(E2) killed four people", E1 and E2 are fully coreferent only when "bombing" is read in its wide sense that denotes the whole attack, not the narrow sense that denotes just the actual exploding of the bomb.

4. **Paraphrase**: One mention is a paraphrase of the other. Here some syntactic differences may occur. Some examples are active/passive transformation ("she gave him the book" / "he was given the book by her"), shifts of perspective that do not add or lose information ("he went to Boston" / "he came to Boston"), etc. No extra semantic information is provided in one mention or the other.

5. **Pronoun**: One mention refers deictically to the DE, as in ("the party" / "that event"), ("the election [went well]" / "it [went well]").

2.2 Quasi-identity

Mention1 is quasi- (partially) identical to mention2 iff they refer to the 'same' DE but one mention includes information that is not contained in the other, not counting information understood/inherited from the context. They are semantically not fully identical, though the core part of the two mentions is. One mention can replace the other, but some information will be changed, added, or lost. (This is the typical case between possible coreferent mentions within a document.)

We distinguish between two core types of partial identity: Membership and Subevent. The essential difference between the two is which aspects of the two events in question differ. Member-of obtains when we have two instances of the same event that differ in some particulars, such as time and location and [some] participants (agents, patients, etc). In contrast, Subevent obtains when we have different events that occur at more or less the same place and time with the same cast of participants.

Membership: Mention1 is a set of similar DEs (multiple instances of the same kind of event), like several birthday parties, and mention2 is one or more of them. More precisely, we say that an event B is a **member** of A if: (i) A is a set of multiple instances of the same type of event (and hence its mention usually pluralized); (ii) B's DE(s) is one or more (but not all) of them; (iii) either or both the time and the place of B's DE(s) and (some of) A's DEs are different. For example, in "I attended three parties(E1) last month. The first one(E2) was the best", E2 is a member of E1. The relation that links the single instance to the set is **member-of**.

Subevent: The DE of mention1 is a script (a stereotypical sequence of events, performed by an agent in pursuit of a given goal, such as eating at a restaurant, executing a bombing, running for election), and mention2 is one of the actions/events executed as part of that script (say, paying the waiter, or detonating the bomb, or making a campaign speech). More precisely, we say that an event B is a **subevent** of an event A if: (i) A is a complex sequence of activities, mostly performed by the same (or compatible) agent; (ii) B is one of

these activities; and (iii) B occurs at the same time and place as A. Here A acts as a kind of collector event. Often, the whole script is named by the key event of the script (for example, in "he planned the explosion", the "explosion" signifies the whole script, including planning, planting the bomb, the detonation, etc.; but the actual detonation event itself can also be called "the explosion"). We call the interpretation of the mention that refers to the whole script its *wide reading*, and the interpretation that refers to just the key subevent the *narrow reading*. It is important not to confuse the two; a wide reading and a narrow reading of a word cannot corefer[3]. The relation that links the narrow reading DE to the wide one is **sub-to**.

Several aspects of the events in question provide key information to differentiate between members and subevents:

1. **Time**: When the time of occurrence of mention1 is temporally 'close enough' to the time of occurrence of mention2, then it is likely that one is a Subevent of the other. More precisely, we say that an event B is a **subevent** of event A if: (i) A and B are both events; (ii) the mentions of A and B both refer to the same overall DE; and (iii) the time of occurrence of B is contained in the time of occurrence of A. But if (i) and (ii) hold but not (iii), and A is a set of events (plural), then B is a member of A. (In (Humphreys et al., 1997), any variation in time automatically results in a decision of non-coreference.)

2. **Space/location**: The location of mention1 is spatially 'close enough' to the location of mention2. More precisely, we say that an event B is a **subevent** of event A if: (i) A and B are both events; (ii) the mentions of A and B both refer to the same overall DE; and (iii) the location of occurrence of B is contained in, or overlaps with, or abuts the location of occurrence of A. But if (i) and (ii) hold but not (iii), and A is a set of events (plural), then B is a member of A.

3. **Event participants**: Mention1 and mention2 refer to the same DE but differ in the overall cast of participants involved. In these cases, the **member** relation obtains, and can be differentiated into subtypes, since participants of events can differ in several ways. For example, if: (i) the mentions of events A and B refer to the same overall DE; and (ii) the participants (agents, patients, etc.) of mention2 are a subset of the participants of mention1, as in "the crowd demonstrated on the square. Susan and Mary were in it", then event B is a **participant-member** of event A. In another example, event B is a **participant-instance-member** of event A if: (i) the mentions of events A and B refer to the same overall DE; and (ii) one or more of the participants (agents, patients, etc.) of mention2 is/are an instance of the participants of mention1, as in "a firebrand addressed the crowd on the square. Joe spoke for an hour", where Joe is the firebrand.

There are other ways in which two mentions may refer to the same DE but differ from one another. Usually these differences are not semantic but reflect an orientation or perspective difference. For example, one mention may include the speaker's evaluation/opinion, while the other is neutral, as in "He sang the silly song. He embarrassed himself", or the spatial orientation of the speaker, as in "she went to New York" / "she came to New York". We treat these cases as fully coreferent.

Sometimes it is very difficult to know whether two mentions are bidirectionally implied, meaning that the two must corefer, or whether they are only quasi-identical (i.e., one entails the other but not vice versa). For example, in "he had a heart attack" / "he died", the two mentions are not identical because one can have a heart attack and not die from it. In contrast, "he had a fatal heart attack" / "he died from a heart attack" are identical. In "she was elected President" / "she took office as President", it is more difficult to decide. Does being elected automatically entail taking office? In some political systems it may, and in others it may not. When in doubt, we treat the case as only quasi-identical. Thus, comparing to examples from Full-Identity: Paraphrase, the following are only quasi-identical because of additional information: "she sold the book" / "she sold Peter the book"; "she sold Peter the book" / "Peter got [*not bought*] the book from her".

[3] For example, in "James perpetrated the shooting. He was arrested for the attack", "shooting" is used in its wide sense and here is coreferent with "attack", since it applies to a whole sequence of events. In contrast, "James perpetrated the shooting. He is the one who actually pulled the trigger", "shooting" is used in its narrow sense to mean just the single act. Typically, a word with two readings can corefer (i.e., be lexically or synonymically identical to) another in the same reading only.

Quasi-identity has been considered in coreference before in (Hasler et al., 2006) but not as extensively, and in (Recasens and Hovy, 2010a; 2011) but applied only to entities. When applied to events, the issue becomes more complex.

3 Two Problems

3.1 Domain and Reporting Events

As described above, inconsistent reporting occurs when a DE stated in reported text contains significant differences from the author's description of the same DE.

To handle such cases we have found it necessary to additionally identify communication events, which we call Reportings, during annotation because they provide a context in which a DE is stated. We identify two principal types of Reporting verbs: locutionary verbs "say", "report", "announce", etc.) and Speech Acts ("condemn", "promise", "support", "blame", etc.). Where the former verbs signal merely a telling, the latter verbs both say and thereby do something. For example in the following paragraph, "admitted" and "say" are communication events:

> Memon **admitted_R.7,in-sayR.3** his involvement in **activities_E.8,in-sayR.3** involving an explosives-laden van near the president's motorcade, police **said_R.3**". Sometimes the same event can participate inside two reporting events, as in "The LA Times **lauded_R.1** the **decision_E.2,in-sayR.1,in-sayR.3**, which the NY Times **lampooned_R.3**.

Though an added annotation burden, the link from a DE to a reporting event allows the analyst or learning system to discount apparent contradictory aspects of the DE and make more accurate identity decisions.

3.2 Unclear Semantics of Events

Sometimes it is difficult to determine the exact relationships between events since their semantics is unclear. In the following, is E.45 coreferent to E.44, or only partially? If so, how?

> Amnesty International has accused both sides of **violating_E.44** international humanitarian law by **targeting_E.45** civilian areas, and ...

We decided that E.44 is not fully coreferent with E.45, since violating is not the same as targeting. Also, E.45 is not a subevent of E.44 since "violating" is not a script with a well-defined series of steps, does not trigger "targeting", and does not occur before "targeting". Rather, targeting is a certain form or example of violation/violating. (It might be easier if the sentence were: "... of **violating** international humanitarian law **by targeting** civilian areas and the human rights group, **by killing** civilians, and **by....**". As such E.45 could be interpreted as a member of E.44, interpreting the latter as a series of violations.)

4 Annotation

To validate these ideas we have been annotating newspaper texts within the context of a large project on automated deep reading of text. This project combines Information Extraction, parsing, and various forms of inference to analyze a small number of texts and to then answer questions about them. The inability of current text analysis engines to handle event coreference has been a stumbling block in the project. By creating a corpus of texts annotated for coreference we are working to enable machine learning systems to learn which features are relevant for coreference and then ultimately to perform such coreference as well.

We are annotating two corpora:

1. The **Intelligence Community (IC) Corpus** contains texts in the Violent Events domain (bombings, killings, wars, etc.). Given the relative scarcity of the partial coreference subtypes, we annotated only instances of full coreference, Subevent, and Member relations. To handle Subevents one needs an unambiguous definition of the scripts in the domain. Fortunately this domain offers a manageable set of events (our event ontology comprises approximately 50 terms) with a subevent structure that is not overly complex but still realistic. We did not find the need to exceed three layers of scriptal granularity, as in *campaign* > {*bombing, attack*} > {*blast, kill, wound*}.

2. The **Biography (Bio) Corpus** contains texts describing the lives of famous people. Typically, these texts are written when the person dies or has some notable achievement. Given the complexities of description of artistic and other creative achievements, we restrict our corpus to achieve-

ments in politics, science, sports, and other more factual endeavors. More important than scriptal granularity in this domain is temporal sequencing.

We obtained and modified a version of the AnCoraPipe entity coreference annotation interface (Bertran et al., 2010) that was kindly given us by the AnCora team at the University of Barcelona. We implemented criteria and an automated method for automatically identifying domain and reporting events. We also created a tool to check and display the results of annotation, and technology to deliver various agreement scores.

Using different sets of annotators (from 3 to 6 people per text), we have completed a corpus of 100 texts in the IC domain and are in process of annotating the Bio corpus. Our various types of full and partial coreference and the associated annotation guidelines were developed and refined over the first third of these documents.

Table 1 shows statistics and inter-annotator agreement for the remaining 65 articles. The average number of domain and reporting events per article is 41.2. We use Fleiss's kappa since we have more than two annotators per article. The (rather low) score for member coreference is not really reliable given the small number of instances.

	Avg no per article	Agreement (Fleiss's kappa)
Full coreference relations	19.5	0.620
Member coreference relations	2.7	0.213
Subevent coreference relations	7.2	0.467

Table 1: Annotation statistics and agreement.

5 Validation and Use

To validate the conceptualization and definitions of full and partial identity relations, we report in (Araki et al., 2013) a study that determines correlations between the Member and Subevent relation instances and a variety of syntactic and lexico-semantic features. The utility of these features to support automated event coreference is reported in the same paper.

We are now developing a flexible recursive procedure that integrates coreference of events and of their pertinent participants (including locations and times). This procedure employs inference in addition to feature-based classification to compensate for the shortcomings of each method alone.

6 Relevant Past Work

The problem of identity has been addressed by scholars since antiquity. In the intensional approach (for example, De Saussure, 1896) a concept is defined as a set of attributes (differentiae), that serve to distinguish it from other concepts; two concepts are identical iff all their attributes and values are. In the extensional approach (Frege, 1982) a concept can be defined as the set of all instances of that concept; two concepts are identical when their two extensional sets are.

Given the impossibility of either approach to support practical work, AI scholars have devoted some attention to so-called Identity Criteria. Guarino (1999) outlines several 'dimensions' along which entities can remain identical or change under transformations; for example, a glass before and after it is crushed is identical with respect to its matter but not its shape; the ACL now and one hundred years hence is (probably) identical as an organization but not in its membership.

There has not been much theoretical work on semantic identity in the NLP community. But there has been a considerable amount of work on the problem of coreference. Focusing on entity coreference are (McCarthy and Lehnert, 1995; Culotta et al., 2007; Ng, 2007; Ng, 2009; Finkel and Manning, 2008; Ng, 2009). Focusing on event coreference are (Humphries et al., 1997; Chen and Zi, 2009; Bejan and Harabagiu, 2008; 2010).

Anaphora and bridging reference are discussed in (Poesio and Artstein, 2005; 2007). Relevant to events is the TIME-ML corpus (Mani and Pustejovsky, 2004; Pustejovsky et al., 2003), which provides a specification notation for events and temporal expressions.

Several corpora contain annotations for entity coreference, including the Prague Dependency Treebank (Kučová and Hajičová. 2004), the ACE corpus (Walker et al., 2006), and OntoNotes (Pradhan et al., 2007).

Most similar to our work is that of (Hasler et al., 2006). In that study, coreferential events and their arguments (also coreference between the arguments) were annotated for the terrorism/security domain, considering five event categories (attack, defend, injure, die, contact), and five event clusters (Bukavu bombing, Peru hostages, Tajikistan hostages, Israel suicide bombing and China-Taiwan

hijacking). They also annotated information about the kind of coreferential link, such as *identity / synonymy / generalization / specialization / other*.

Our work takes further the ideas of (Hasler et al., 2006) and (Recasens et al., 2011) in elaborating the types of full and partial identity, as they are manifest in event coreference.

7 Conclusion

The problem of entity and event identity, and hence coreference, is challenging. We provide a definition of identity and two principal types of quasi-identity, with differentiation based on differences in location, time, and participants. We hope that these ideas help to clarify the problem and improve inter-annotator agreement.

Acknowledgments

Our grateful thanks goes to Prof. Antonia Martí and her team for their extensive work on the modifications of the AnCoraPipe annotation interface.

References

Araki, J., T, Mitamura, and E.H. Hovy. 2013. Identity and Quasi-Identity Relations for Event Coreference. Unpublished manuscript.

Bejan, C.A. and S. Harabagiu. 2008. A Linguistic Resource for Discovering Event Structures and Resolving Event Coreference. *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 08)*.

Bejan, C.A. and S. Harabagiu. 2010. Unsupervised Event Coreference Resolution with Rich Linguistic Features. *Proceedings of the 48th conference of the Association for Computational Linguistics (ACL 10)*.

Bertran, M., O. Borrega, M.A. Martí, and M. Taulé, 2010. AnCoraPipe: A New Tool for Corpora Annotation. Working paper 1: TEXT-MESS 2.0 (Text-Knowledge 2.0). Available at http://clic.ub.edu/files/AnCoraPipe_0.pdf

Chen, Z. and H. Ji. 2009. Graph-based Event Coreference Resolution. *Proceedings of the ACL-IJCNLP 09 workshop on TextGraphs-4: Graph-based Methods for Natural Language Processing*.

Culotta, A., M. Wick, and A. McCallum. 2007. First-order probabilistic models for coreference resolution. *Proceedings of the HLT/NAACL conference*.

De Saussure, F. 1896. *Course in General Linguistics*. Open Court Classics.

Finkel, J.R. and C.D. Manning. 2008. Enforcing transitivity in coreference resolution. *Proceedings of the ACL-HLT conference,* pp. 45–48.

Florian, R., J F Pitrelli, S Roukos, I Zitouni. 2010. Improving Mention Detection Robustness to Noisy Input. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Frege, G. 1892. On Sense and Reference. Reprinted in P. Geach and M. Black (eds.) *Translations from the Philosophical Writings of Gottlob Frege.* Oxford: Blackwell, 1960.

Guarino, N. 1999. The Role of Identity Conditions in Ontology Design. In C. Freksa and D.M. Mark (eds.), *Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science*. *Proceedings of International Conference COSIT '99*. Springer Verlag.

Hasler, L., C. Orasan, and K. Naumann. 2006. NPs for Events: Experiments in Coreference Annotation. *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC-06)*, pp. 1167–1172.

Hasler, L. and C. Orasan. 2009. Do Coreferential Arguments make Event Mentions Coreferential? *Proceedings of the 7th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC 09)*, pp 151–163.

Humphreys, K., R. Gaizauskas and S. Azzam. 1997. Event Coreference for Information Extraction. *Proceedings of the ACL conference Workshop on Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts (ANARESOLUTION 97)*.

Kučová, L. and E. Hajičová. 2004. Coreferential relations in the Prague Dependency Treebank. *Proceedings of the DAARC workshop*, pp. 97–102.

Mani, I. and J. Pustejovsky. 2004. Temporal Discourse Models for Narrative Structure. *Proceedings of the ACL 2004 Workshop on Discourse Annotation*.

McCarthy, J.F. and W. Lehnert. 1995. Using Decision trees for Coreference Resolution. *Proceedings of the IJCAI conference*.

Mill, J.S. 1872. *A System of Logic,* definitive 8th edition. 1949 reprint, London: Longmans, Green and Company.

Ng, V. 2007. Shallow Semantics for Coreference Resolution. *Proceedings of the IJCAI conference*.

Ng, V. 2009. Graph-cut-based Anaphoricity Determination for Coreference Resolution. *Proceedings of the NAACL-HLT conference*, pp. 575–583.

Poesio, M. and R. Artstein. 2005. The reliability of anaphoric annotation, reconsidered: Taking ambiguity into account. *Proceedings of the ACL Workshop on Frontiers in Corpus Annotation II*.

Poesio, M. and R. Artstein. 2008. Anaphoric annotation in the ARRAU corpus. *Proceedings of the LREC conference*.

Pradhan, S., E.H. Hovy, M. Marcus, M. Palmer, L. Ramshaw, and R. Weischedel 2007. OntoNotes: A Unified Relational Semantic Representation. *International Journal of Semantic Computing* 1(4), pp. 405–420.

Pustejovsky, J., J. Castaño, R. Ingria, R. Saurí, R. Gaizauskas, A. Setzer and G. Katz. 2003. TimeML: Robust Specification of Event and Temporal Expressions in Text. *Proceedings of IWCS-5, Fifth International Workshop on Computational Semantics*.

Recasens, M. and E.H. Hovy. 2010a. Coreference Resolution across Corpora: Languages, Coding Schemes, and Preprocessing Information. *Proceedings of the Association of Computational Linguistics conference (ACL 10)*.

Recasens, M. and E.H. Hovy. 2010b. BLANC: Implementing the Rand Index for Coreference Evaluation. *Journal of Natural Language Engineering* 16(5).

Recasens, M., E.H. Hovy, and M.A. Martí. 2011. Identity, Non-identity, and Near-identity: Addressing the Complexity of Coreference. *Lingua*.

Taulé, M., M.A. Martí. and M. Recasens. 2008. AnCora: Multilevel Annotated Corpora for Catalan and Spanish. *Proceedings of the LREC 08 conference*, pp. 96–101.

Walker, C., S. Strassel, J. Medero 2006. The ACE 2005 multilingual training corpus. Linguistic Data Consortium, University of Pennsylvania, Philadelphia.