Cross-Document, Cross-Language Event Coreference Annotation Using Event Hoppers

Zhiyi Song, Ann Bies, Justin Mott, Xuansong Li, Stephanie Strassel, Christopher Caruso
Linguistic Data Consortium
Suite 810, 3600 Market Street, Philadelphia, Pennsylvania, 19104, USA
{zhiyi, bies, jmott, xuansong, strassel, ccaruso}@ldc.upenn.edu

Abstract

We discuss the development and implementation of an approach for cross-document, cross-lingual event coreference for the DEFT Rich Entities, Relations and Events (Rich ERE) annotation task. Rich ERE defined the notion of event hoppers to enable intuitive within-document coreference for the DEFT event ontology, and the expansion of coreference to cross-document, cross-lingual event mentions relies crucially on this same construct. We created new annotation guidelines, data processes and user interfaces to enable annotation of 505 documents in three languages selected from data already labeled for Rich ERE, yielding 389 cross-document event hoppers. We discuss the data creation process and the central role of event hoppers in making cross-document, cross-lingual coreference decisions. We present the challenges encountered during annotation along with three directions for future work.

Keywords: Events, Information Extraction, Coreference, Linguistic Resources

1. Introduction

This paper discusses the development and implementation of annotation for cross-document and cross-lingual event coreference as an expansion to the Rich Entities, Relations, and Events (ERE) annotation task first defined as part of DARPA's Deep Exploration and Filtering of Text (DEFT) program (DARPA, 2012). The goal of the DEFT program is to develop technologies capable of extracting knowledge from unstructured text in multiple languages and genres. ERE annotation was developed at Linguistic Data Consortium (LDC) to support multiple research directions and evaluations in DEFT. ERE builds on the approach to labeling entities, relations, events and their attributes under a pre-defined taxonomy, following the approach used in Automatic Content Extraction (ACE) (LDC, 2005; Walker et al., 2006; Song et al., 2015; Mott et al., 2016).

One important DEFT use case is automatically building a structured Knowledge Base (KB) from scratch. This task, known as Cold Start, is one of several tasks relevant to DEFT that are evaluated in the NIST Knowledge Base Population evaluation series (NIST 2017; Getman et al, 2018). Given DEFT's focus on whole-corpus understanding culminating in the Cold Start task, Rich ERE annotation has evolved over the course of the program to emphasize cross-document and cross-lingual approaches. Rich ERE event annotation includes 9 event types and 38 subtypes (e.g. Conflict.Attack, Contact.Meet, Movement.TransportPerson). For each event mention annotators label the most salient word evoking the event (the "event trigger"), the event type and subtype, the realis status (Actual, Other or Generic), all of the event's arguments (e.g. agent, instrument) and several attributes like temporal information (Aguilar et al., 2014; Song et al., 2015). Multiple mentions of the same event in the document are labeled for coreference, utilizing the notion of event hoppers. While stricter approaches to event coreference typically require all event features (including arguments and attributes) to be identical, event hoppers relax this requirement to enable coreference of two events that are intuitively the same although certain features may differ. For instance, two reports about the same terrorist incident may differ with respect to the number of perpetrators, especially in the immediate aftermath of the event when facts are still being uncovered. The event hopper approach allows for coreference of such event mentions where other event coreference schemas may not. While event hoppers were originally introduced to support within-document Rich ERE annotation, extending them to cross-document and cross-lingual event coreference was a natural progression since the approach results in a more complete Knowledge Base, particularly when the KB reflects information extracted from multiple languages, sources and genres. In the sections that follow we present the results of our effort to define an approach to cross-document, cross-lingual event coreference using event hoppers as part of the Rich ERE annotation task in DEFT.

2. Related Work

There have been other efforts that have captured some variety of cross-document event-event coreference, which have informed our design of the cross-document event coreference annotation task. These include topic-clustering of documents and pair-wise comparison of event mentions.

The EventCorefBank (ECB) (Bejan and Harabagiu, 2008) contains 482 documents clustered into 43 topics that were annotated for within-document and cross-document coreference according to the TimeML specification (Pustejovsky et al., 2013). Lee et al. (2012) extended ECB annotation, following the OntoNotes guidelines (Pradhan et al., 2007). These studies required both matching predicates and matching arguments for event coreference. The ECB+ corpus (Cybluska and Vossen, 2014a) is an extension of ECB with the addition of source documents as well as event components, using the CROMER (CROss-document Main Events and entities Recognition) tool (Girardi et al., 2014). Event coreference in ECB+ required matching time, place, and participants (Cybluska and Vossen, 2014b).

A richer event-event relation annotation scheme (Hong et al., 2016) was developed to capture cross-document event-event relations, including coreference. This data was constructed using ACE 2005 data and supplemented with data collected by researchers; event relations including coreference were annotated by pairwise comparison of events from documents within a given topic. Event coreference required event arguments in the pair to match.

To support the cross-document component of Event Argument evaluation in TAC-KBP 2016 (NIST, 2016),

LDC manually identified strings of text that contained the event corresponding to a series of manually curated queries (Ellis et al., 2016). The evaluated systems, however, failed to produce any entry points for those queries.

However, we are not aware of previous attempts at labeling cross-document coreference using the event hopper framework, which focuses on annotators' intuitions about event reference and allows event argument mismatches as well as event mention realis mismatches. This approach allows annotators to group together more event mentions and arguments than previous stricter coreference approaches, on both the within-document and cross-document level.

3. Annotation Approach

Cross-document and cross-lingual event coreference annotation for the current effort starts with within-document Rich ERE event hopper annotation as input, and coreferences event hoppers from different documents or languages. In Rich ERE event annotation, every tagged event mention is put into an event hopper. All event mentions that refer to the same event occurrence are grouped into the same event hopper, with the result that each event hopper consists of one or more event mentions. The criteria for judging whether hoppers are coreferential or not are the same as those outlined in the description of the event hoppers in Song et al. (2015):

- They are intuitively the same event
- They have the same event type and subtype
- Temporal and place arguments don't need to match, but need to be the same general scope
- Event arguments may be non-coreferential or conflicting
- Realis status may be different

The event hopper concept is a more inclusive, less strict notion of event coreference than that used in ACE and other schemes, handling within-document event coreference using the notion of event hoppers (as above) which permits intuitive coreference and allows non-matching arguments or realis (see the following subsections for examples). Cross-document coreference is inherently more difficult for annotators than within-document, because the cross-document and cross-language aspect of the task requires annotators to fully digest multiple documents being compared and to develop an understanding of the overall topic as context. When comparing pairs of event hoppers from individual documents, it can be difficult to understand how the individual hoppers fit into the larger picture of the topic as a whole. Moving from annotating within-document event hoppers to cross-document coreference highlights several points where the less strict event hopper concept is necessary, such as differing granularity or realis, and multiple occurrences.

3.1 Event Argument Granularity

Perhaps the most common reason that event mentions are not strictly coreferential is event arguments that differ in granularity or sometimes conflict. For example,

> S1: **Attack** in Baghdad on Thursday
>
> S2: **Bombing** in the Green Zone last week

The journalistic nature of newswire documents lends itself to repetitive language that varies in levels of specificity. Annotators would that these two Conflict.Attack events are coreferential when looking at the context of the documents. Annotators should treat two event hoppers as coreferential even when their arguments are not identical, if the events are intuitively the same.

In the cross-document task, it can be particularly difficult to decide whether the same event is being reported with differing arguments over time, or if it is a different event.

> S1: Policía china **detiene** a 118 sospechosos en un caso de contaminación ambiental.
> *Chinese police are arresting 118 suspects in a case of environmental pollution.*
>
> S2: Ocho personas han sido **detenidas** por haber realizado vertidos ilegales en un río local de la provincia suroccidental china de Yunnan.
> *Eight people have been arrested for illegally dumping in a local river in the southwestern Chinese province of Yunnan.*

Here, it is possible that the 118 suspects mentioned in the first sentence are a superset containing the 8 suspects in the second sentence from a different document. In such cases, the context of the entire document is taken into account. If doing so still does not resolve the ambiguity, then the default is to not coreference the events.

3.2 Factually Conflicting Event Arguments

Event hopper annotation also addresses factually conflicting Event Arguments. Unstructured text such as discussion forum documents, commonly contains misinformation or conflicting assertions. Newswire data covering breaking events may also present seemingly conflicting reports based on what it known at the time of publication. Annotators need to use general information or real knowledge or their own judgment for these coreference decisions.

In the following example, the first document contains an assertion, but in the second document, the information relayed contradicts that in the first document:

> S1: John Smith was **killed** in Canada.
>
> S2: Many people believe that the man was actually shot **dead** in Iowa.

These two event mentions have conflicting place argument (Canada vs. Iowa), but they are interpreted as coreferential, because both mentions refer to the Life.Die event of "John Smith" (also mentioned as "the man").

3.3 Event Hoppers with Different Realis State

Cross-document event coreference annotation includes realis states from the Rich ERE event annotation. A future or hypothetical event mention would be marked "Other", while a similar event mention that refers to an asserted event would be marked "Actual". Different temporal states of the same event should be coreferred. For instance,

S1: Barack Obama will **visit** [OTHER] London next month, the White House has confirmed.

S2: US President Barack Obama **arrived** [ACTUAL] at the Globe Theatre on April 23, 2016 in London, England

When approaching two mentions in cross-document cross-lingual event coreference annotation, "Other" and "Actual" event hoppers can be coreferred if they are intuitively the same event. The event hopper pair in the above example would be judged as coreferential.

3.4 Multiple Occurrences of Events of the Same Type and Subtype

Multiple instances of the same type/subtype of event can occur within documents or within a topic. For example,

S1: Angela Merkel was first **elected** in 2005 and re-elected in 2009 and 2013.

S2: Merkel's re-**election** was never in doubt.

The first sentence contains three separate Personnel.Elect events; it is easy to distinguish them because of the explicit mention of time arguments. However, the second sentence could refer to either 2009 or 2013. The broader context of the document can be used to resolve this ambiguity. However, if it is not possible to resolve in context, then the mentions will not be coreferred.

3.5 Event Hoppers with Conflicting Volition

Event mentions can be coreferred when the arguments of the event are referred to with differing levels of volition in different documents. For example,

S1: The defendant was **dragged** kicking and screaming from the courtroom.

S2: The defendant **left** the courtroom.

Here, the two Movement.TransportPerson event mentions can be included in the same hopper, even though in the first instance the it is clear the entity argument is being transported involuntarily, whereas in the second the same argument seems to be a voluntary participant.

4. Methodology

To perform the cross-document event hopper coreference task, annotators were presented with a pair of event hoppers from two different documents to compare and judge as coreferential or not. The presented event hoppers already included all event mentions that were judged as coreferential from within-document annotation (as well as event hoppers containing only singleton event mentions).

4.1 Source Data

The data consisted of 505 "core" source documents which were annotated with Rich ERE for the TAC KBP 2016 evaluations (Ellis et al., 2016). All of these documents (254 Newswire (NW), and 251 Discussion Forum (DF) threads) were manually selected using a topic-driven approach to ensure appropriate coverage of event types, ambiguous entities, and entities referenced only by nominal mentions (Ellis et al., 2016). Table 1 shows the document and token count distribution by language and genre for the TAC KBP 2016 evaluation "core" set.

Language	Genre	Documents	Tokens
Chinese	NW	85	43,338
Chinese	DF	82	78,675
English	NW	85	41,622
English	DF	84	46,282
Spanish	NW	84	26,228
Spanish	DF	85	40,703
Total	--	505	276,848

Table 1: TAC KBP 2016 "core" set

A wide variety of newsworthy topics were included, with a total of 61 topics, such as: South China Sea Tension, Spanish Train Derailment, Syrian Revolution, Typhoon Haiyan, etc. Some topics included documents in two languages, some in only one language, and ten of the topics included documents from all three languages.

ERE annotation for the "core" set followed the Rich ERE annotation guidelines (Song et al., 2015), with the exception that the inventory of Rich ERE event types and subtypes was reduced to 18 types and subtypes.

4.2 Data Pruning

Exhaustive manual pairwise comparison of all event hoppers in the corpus for coreference was infeasible. We therefore limited the annotation scope by (1) topic, (2) event type/subtype, and (3) realis, in order to make the annotation task more tractable.

Within each topic we identified documents with event hoppers of the same type and subtype. Within that pool we then selected a "seed hopper", while hoppers from all other documents in the pool became "candidate hoppers". Annotators judged each candidate hopper against the seed hopper and made a coreference decision. Candidate hoppers that were not co-referred with the current seed hopper then became available as candidates or seeds for the next iteration of coreference. The iterations continued until the annotators had judged all hoppers in the pool against one another.

Table 2 shows the total number of pools, event hoppers and hopper pairs for cross-document event coreference annotation. The actual number of pairs annotated is smaller than the total number of event hopper pairs due to the exclusion of the non-seed hoppers from the seed document in each iteration and hoppers that were judged as coreferential from previous iterations.

Language	Pools	Hoppers	Max. hopper pairs	Annotated pairs
Chinese	140	1643	21761	14217
English	166	2454	49124	27946
Spanish	143	1234	11613	7265
Total	449	5331	82398	49428

Table 2: Pool, hopper and hopper pair counts for cross-document annotation

4.3 Cross-lingual Annotation

After cross-document coreference annotation within each language was completed, we clustered the coreferential event hoppers into cross-document event hopper clusters.

An event hopper cluster contains two or more event hoppers from different documents. These event hopper clusters were candidates for cross-lingual linking annotation. Event hoppers that were not included in event hopper clusters were considered as singleton event hoppers and not fed into cross-lingual linking annotation. We pivoted cross-lingual coreference though English by linking Chinese and Spanish hopper clusters to English hopper clusters. No additional direct linking of Chinese and Spanish hoppers was performed due to the scarcity of Chinese-Spanish bilingual specialists. Clusters were linked if they were

- From documents of the same topic
- The same event type and subtype
- Coreferential according to event hopper criteria

Table 3 shows the total number of pools, event hoppers and hopper pairs for cross-lingual event coreference annotation.

Language	Pools	Hopper clusters	Annotated pairs
Cmn_eng	29	102	237
Spa_eng	25	87	217
Total	54	189	454

Table 3: Pool, hopper and hopper pair counts for cross-lingual annotation

5. Results and Challenges

This cross-document annotation effort resulted in 892 coreference pairs and 389 event hopper clusters, as shown in Table 3.

Language	Total pairs	Coreferen -tial pairs	Total hoppers	Hopper clusters
Chinese	14217	256	1643	108
English	33527	423	2454	195
Spanish	7265	213	1234	86
Total	55009	892	5329	389

Table 4: Annotation results for cross-document coreference annotation

Out of a total of 55,009 hopper pairs, only 892 pairs were judged as coreferential, with a coreference ratio of 1.6%. Certain event types and subtypes have a higher coreference ratio – for example, the subtypes of the Personnel type: Elect 32%, StartPosition 16%, and EndPosition 24%.

The coreferenced event hoppers resulted in a total of 389 hopper clusters, with most hopper clusters containing 2-3 coreferenced event hoppers. The largest hopper cluster in Chinese had 15 coreferenced event hoppers, in Spanish 12, and in English 16.

The cross-lingual event hopper linking annotation based on the 189 cross-document hoppers resulted in a total of 28 hopper clusters, with 13 tri-lingual clusters and 15 bi-lingual clusters.

Although we selected a dataset of documents sharing the same topics to increase the chance of cross-document event hopper coreference in the data, there were still relatively few coreferenced event hoppers in this data. We took advantage of an existing topic-annotated corpus for this work, but the topics were not originally designed for this coreference task. A more suitable corpus for future work would have more specifically targeted topics that are more aligned to the event ontology that we are working with. The Personnel type had a high ratio of coreference in this corpus in part because this particular event type aligned well with the pre-existing topics (e.g., Angela Merkel's Third Electoral Win, Presidential Election of Nicholas Maduro, Taiwan Presidential Election).

Resource limitations did not permit dual annotation and calculation of inter-annotator agreement numbers during the initial pilot, but this is a necessary part of any future work.

5.1 Annotation Efficiency

During annotation, annotators judged one pair of event hoppers at a time. Some of the pools had many more event hoppers than others, which resulted in a substantial number of pairs to be judged. Such pools can also have a very long tail of judgments to be made through the iterations, and this is time consuming for annotation. Figure 2 illustrates the long tails of some event hopper pools. The reason that

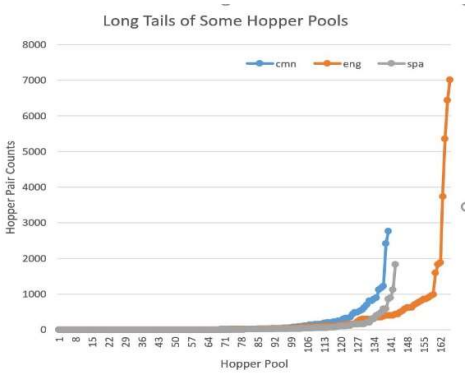

Figure 2: Some hopper pools have very long tails

English had more pools with long tails is that the English pools tend to be larger in this corpus, the largest being the Conflict.Attack pools for the "Syrian Revolution" and "Egypt Protests" topics, with 114 and 119 event hoppers.

5.2 Aggregate Events and Subevents

Drawing the distinction between subevents of aggregate events and events that are intuitively the same but with arguments of differing granularity remains a challenge for annotation, more so for cross-lingual annotation. An aggregate event is an event that may be composed of two or more subevents, which are only partially coreferential. Aggregated events and subevents have a parent-child relation, and the subevents themselves have a sister-like relation (Araki et al., 2014).

In the example below, event hopper 1 is the aggregate of protests (Conflict.Demonstration) occurring throughout the country. Each of the subsequent hoppers refers to a separate subevent of the aggregate event in eh1.

> S1: **Protests** (eh1) broke out throughout Bolivia over an increase in gas prices.
>
> S2: The **march** (eh2) in the capital began peacefully but clashes with police erupted near the main plaza where the government palace is located.

S3: Taxi drivers held a **strike** (eh3) that largely paralyzed La Paz on Thursday.

Often the challenge of aggregated events and subevents arises when the same aggregate event occurs at different times or places. For example, the Conflict.Demonstrate events in S1, S2 and S3 are co-referred, as they occur in about the same time period and same place, but the event in S4 doesn't belong to this event hopper cluster as it happened at a different time, although it was a subevent of the same aggregate event "protest in Egypt".

S1: Morsi's supporters, who have been holding sit-ins and **demonstrations** (eh1) since the president's ouster. (2013-07-16)

S2: 穆尔西和穆兄会的支持者在多座城市发起大规模**游行** (eh2). (2013-07-19)
*Supporters of Morsi and the Muslim Brotherhood launched big **protests** in multiple cities.*

S3: 土耳其民众举行集会声援埃及**示威** (eh3)者 (2013-08-17)
*People in Turkey gathered to support Egyptian **protesters**.*

S4: 当天，在开罗、亚历山大、苏伊士、法尤姆等省份爆发了不同规模的**示威游行** (eh4) (2013-12-14)
*On that same day, **protests** erupted in Cairo, Alexandria, Suiz, Faiyum and other provinces.*

Annotators should not place the aggregate event in the same hopper as any of its subevents, and likewise should not place the subevents in a hopper with each other. However, it is sometimes difficult to distinguish subevents as opposed to the argument granularity examples in Section 3.1, and this is an area that we will continue to investigate.

6. Future Directions

The work reported here has suggested three directions for future work: First, the annotation pipeline for cross-document/cross-lingual event coreference could be further optimized by leveraging existing entity linking annotation of event arguments (Ellis et al., 2016). It is expected that event hoppers sharing arguments that are linked to the same node of an entity knowledge base would have a greater likelihood of being coreferent.

Second, developing a KB of events, or events that occur within each topic, would allow document-level event hoppers to be linked to the KB. Similar to work that has been done for Entity Detection and Linking (EDL) (Ji et al., 2010), linking to such a KB of events would reduce the need to compare every relevant document-level event hopper to every potentially coreferent hopper, since many document-level hoppers could be linked directly to the event KB. The remaining document-level event hoppers that are not found in the KB would still need to be coreferenced via a pairwise comparison as in this paper (as NIL clusters are created for EDL entities). This direction, however, would require building such a KB before annotation. Giraldi et al. (2014) adopted this approach and demonstrated feasibility, but the event and coreference definitions used were quite different from the ERE framework.

Third, the cross-document/cross-lingual coreferenced event hoppers that are the result of this process can now be used in other event-event relations, such as part-whole, causation, or event sequencing. Using corpus-wide coreferenced event hoppers (rather than individual event mentions) as the arguments of event-event relations would allow for a corpus-wide view of events and event relations, which is critical for corpus-wide evaluation and understanding (Hong et al., 2016).

7. Conclusion

We created a small corpus annotated for cross-document and cross-lingual event coreference in 505 documents in three languages. Although we leveraged existing ERE annotation as input, this task required the development of new annotation guidelines, new data processes and user interfaces, and the creation of new cross-document and cross-lingual annotation. The more intuitive, coarser grained event hopper concept that was originally developed as part of within-document Rich ERE annotation (Song et al., 2015) has proven to be well suited for the type of event coreference that is possible across documents and across languages. We plan to continue with corpus-wide event coreference using the event hopper concept.

Although this corpus is relatively small so far, it does provide data in support of developing a corpus-wide understanding of events and the entities participating in those events. The corpus includes both positive examples of corpus-wide event hopper coreference, both cross-document and cross-lingual, and also negative coreference judgements of many more potential event pairs. Positive and negative judgements both provide useful training data for identifying corpus-wide events and event argument clusters.

The cross-document event coreference annotation described in this paper has been distributed to DEFT performers and to participants in the 2017 NIST TAC KBP evaluations. Following its use in these evaluations, the data and annotations will be published in LDC's public catalog, making the resource broadly available for language-related research, education and technology development.

8. Acknowledgements

This material is based on research sponsored by Air Force Research Laboratory and Defense Advanced Research Projects Agency under agreement number FA8750-13-2-0045. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory and Defense Advanced Research Projects Agency or the U.S. Government.

We also thank Jonathan Wright for building the cross-document annotation tool that makes the annotation possible.

9. Bibliographical References

Aguilar, J., Beller, C., McNamee, P., Van Durme, B., Strassel, S., Song, Z., & Ellis, J. (2014). A comparison of the events and relations across ACE, ERE, TAC-KBP, and FrameNet annotation standards. In Proceedings of the Second Workshop on EVENTS: Definition, Detection, Coreference, and Representation (pp. 45-53).

Araki, J., Liu, Z., Hovy, E. H., & Mitamura, T. (2014). Detecting Subevent Structure for Event Coreference Resolution. In Proceedings of LREC 2014.

Bejan, AC. and Harabagiu, S. (2008). A Linguistic Resource for Discovering Event Structures and Resolving Event Coreference. In Proceedings of LREC 2008.

Bies, A., Song, Z., Getman, J., Ellis, J., Mott, J., Strassel, S., Palmer, M., Mitamura, T., Freedman, M., Ji, H., and O'Gorman, T. (2016). A Comparison of Event Representations in DEFT. In Proceedings of the Fourth Workshop on Events (pp. 27-36).

Cybulska, A., & Vossen, P. (2014a). Using a sledgehammer to crack a nut? Lexical diversity and event coreference resolution. In Proceedings of LREC 2014.

Cybulska, A., & Vossen, P. (2014b). Guidelines for ECB+ annotation of events and their coreference. Technical Report NWR-2014-1, VU University Amsterdam.

DARPA. 2012. *Broad Agency Announcement: Deep Exploration and Filtering of Text (DEFT)*. Defense Advanced Research Projects Agency, DARPA-BAA-12-47.

Ellis, J., Getman, J., Fore, D., Kuster, N., Song, Z., Bies, A., & Strassel, S. (2016). Overview of Linguistic Resources for the TAC KBP 2016 Evaluations: Methodologies and Results. In Proceedings of TAC KBP 2016 Workshop, National Institute of Standards and Technology.

Getman, J., Ellis, J., Strassel, S., Song, Z., & Tracey, J. (2018). Laying the Groundwork for Knowledge Base Population: Nine Years of Linguistic Resources for TAC KBP. In Proceedings of LREC 2018.

Girardi, C., Speranza, M., Sprugnoli, R., & Tonelli, S. (2014). CROMER: a Tool for Cross-Document Event and Entity Coreference. In Proceedings of LREC 2016.

Hong, Y., Zhang, T., O'Gorman, T., Horowit-Hendler, S., Ji, H., & Palmer, M. (2016). Building a Cross-document Event-Event Relation Corpus. LAW X, 1.

Ji, H., Grishman, R., Dang, H. T., Griffitt, K., & Ellis, J. (2010). Overview of the TAC 2010 Knowledge Base Population Track. In Third Text Analysis Conference (TAC 2010).

Lee, H., Recasens, M., Chang, A., Surdeanu, M., & Jurafsky, D. (2012). Joint entity and event coreference resolution across documents. In Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (pp. 489-500). Association for Computational Linguistics.

Linguistic Data Consortium. (2005). *ACE (Automatic Content Extraction) English Annotation Guidelines for Events Version 5.4.3*.

Mott, J., Bies, A., Song, Z., & Strassel, S. (2016). Parallel Chinese-English Entities, Relations and Events Corpora. In Proceedings of LREC 2016.

NIST. (2016). TAC 2016 Cold Start KB Track. https://tac.nist.gov/2016/KBP/ColdStart/index.html. March 2017. Accessed September, 2017.

NIST. (2017). TAC 2017 Cold Start KB Track. https://tac.nist.gov/2017/KBP/ColdStart/index.html. May 2017. Accessed September, 2017.

Pradhan, S. S., Ramshaw, L., Weischedel, R., MacBride, J., & Micciulla, L. (2007). Unrestricted coreference: Identifying entities and events in OntoNotes. In Semantic Computing, 2007. ICSC 2007. International Conference on (pp. 446-453). IEEE.

Pustejovsky, J., Castano, J.M., Ingria, R., Sauri, R., Gaizauskas, R.J., Setzer, A., Katz, G. and Radev, D.R. (2003). TimeML: Robust specification of event and temporal expressions in text. New directions in question answering, 3, 28-34.

Song, Z., Bies, A., Strassel, S., Riese, T., Mott, J., Ellis, J., Wright, J., Kulick, S., Ryant, N., and Ma, X. (2015). From Light to Rich ERE: Annotation of Entities, Relations, and Events. In Proceedings of the 3rd Workshop on Events at the NAACL-HLT.

10. Language Resource References

Walker, C., Strassel, S., Medero, J., and Maeda, K. (2006). *ACE 2005 Multilingual Training Corpus*. Linguistic Data Consortium, LDC Catalog No.: LDC2006T06.