

Distant annotation of Chinese tense and modality

Nianwen Xue, Yuchen Zhang and Yaqin Yang

Computer Science Department

Brandeis University, Waltham, USA

{xuen, yuchenz, yaqin}@brandeis.edu

Abstract

In this paper we describe a “distant annotation” method by which we mark up tense and modality of Chinese eventualities via a word-aligned parallel corpus. We first map Chinese verbs to their English counterpart via word alignment, and then annotate the resulting English text spans with coarse-grained tense and modality categories that we believe apply to both English and Chinese. Because English has richer morpho-syntactic indicators for tense and modality than Chinese, we hope this distant annotation approach will yield more consistent annotation than if we annotate the Chinese side directly. We report experimental results that show this expectation is largely borne out.

1 Introduction

It is often the case that linguistic information that is hidden in one language is directly observable in another. This is particularly true for typologically distant language pairs. For example, tense and number are “invisible” in languages like Chinese but they are explicitly marked in languages like English where there is a more developed morphological system. In English, for example, tense is grammaticalized as an inflectional morpheme attached to a verb. In Chinese, on the other hand, such morphological cues are rare or non-existent, and the underlying semantic tense has to be inferred from the context. (1) is an example Chinese sentence and its verb 举行 (“hold”) has no morphological inflections of tense. However, a notional semantic tense that indicates the temporal location of the event denoted

by 举行 can be inferred by the time expression 明年 (“next year”).

- (1) 大会 明年 在新加坡 举行。
conference next year in Singapore hold .

“The conference will be held in Singapore next year.”

In this paper, we describe a method of annotating the tense of a Chinese sentence by annotating the tense of its English translation and then projecting this annotation back onto the Chinese sentence. Specifically, we identify all English text spans that are aligned to a Chinese verb in a word-aligned parallel Chinese-English corpus. Then all the English text spans will be annotated with tense and modality. Note that the resulting English text spans after such mapping may not necessarily be English verbs because a Chinese verb may be translated into an English noun, or words of other parts-of-speech. In (2), for example, “appointment” is translated into a Chinese verb “赴约”. Nevertheless, such English text spans can still be treated as “anchors” of tense and modality. Our hypothesis is that we are more likely to obtain consistent annotation by annotating the English translation rather than the Chinese source directly because the morpho-syntactic cues in English are good indicators of tense and they constrain the choices that an annotator has to make during the annotation process.

- (2) a. One day in August 2005 , Abather was driving Balqes , six months pregnant with their second child , to a doctor ’s appointment .

- b. 2005年8月的一天，阿巴舍正开车送怀着他们第二个孩子已有六个月的巴尔科斯赴约去医生处。

It is important to note that the target of our annotation is the underlying semantic tense, not the grammatical tense. The semantic tense can be interpreted as a relation between the time that an event occurs and a reference time in the sense of (Reichenbach, 1947). In a written text, this reference time is usually the time when the document is created or the time when another event occurs. This relation does not change regardless of whether it has a morpho-syntactic manifestation when a sentence is translated into a different language.

The grammatical tense, indicated by the morpho-syntactic cues, do not always have a one-to-one correspondence with the semantic tense. In English, it has long been noted that the same surface morpho-syntactic form may indicate very different underlying semantic notions of tense (Dinsmore, 1991; Cutrer, 1994; Fauconnier, 2007). For example, the morpho-syntactic form “are playing” in both (3a) and (3b) indicates present progressive tense, but underlyingly, while in (3a), the present tense indeed indicates a *present* time, in (3b), the present tense indicates a *future* time. Typically, however, these morpho-syntactic cues have a relatively straightforward correspondence with their underlying semantics in the source language, making the manual annotation of the underlying semantics very straightforward.

- (3) a. They are playing soccer in the park.
b. They are playing soccer next Tuesday.

In addition to the “polysemous” finite tense forms, there are also a large proportion of English verbs that appear in non-finite forms, which can also be construed as having multiple semantic tense values. The examples in (4) show that infinitives and gerunds can indicate past (4a, 4b), present (4c, 4d) or future (4e, 4f) times. When projected onto a target language, these non-finite forms need to be disambiguated so that the correct semantic tense can be accurately inferred from the context in the target language.

- (4) a. I enjoyed reading the book.

- b. The prisoner managed to escape.
c. I enjoy reading.
d. To err is human, to forgive is divine.
e. The prisoner is considering escaping.
f. The prisoner hopes to escape tomorrow.

Tense is intricately related to modality. Although both (3b) and (4f) indicate future times, they differ in that (3b) indicates an event that has been previously scheduled to happen while (4f) describes an intended event that may or may not materialize. There is a higher level of certainty that the soccer playing event is going to happen than that escaping event will. In other words, although the tense is the same, the modality, which accounts for the degree of certainty of an event, is different.

The rest of the paper is organized as follows. In Section 2, we describe our annotation framework. We present the results of a preliminary annotation experiment in Section 3. Section 4 describes related work and Section 5 concludes the paper.

2 Specifications

As described in Section 1, each Chinese verb instance is mapped to a text span in English and then annotation is performed on English text by labeling these text spans with tense and modality categories. Each text span is annotated along three dimensions to support the planned automatic inference of tense and modality on the Chinese side. The first dimension is the *semantic tense*, and the annotator must indicate whether the text span describes a past, present, or future event state. The second dimension is *event type* that indicates whether the text span represents a habitual event, an ongoing event, a completed event, an episodic event, or a state. The event type is annotated because it has been shown (Smith, 2001; Smith and Erbaugh, 2005; Xue, 2008) that in a language without explicit tense markers, event types are good indicators of tense. The third dimension is *modality*. The modality dimension is broadly construed and it classifies events or states as actual, intended (which encompasses expected, planned events), hypothetical (as in conditional clauses) or modalized. An event or state is modalized if it occurs with a modal verb that

indicates possibility, necessity, or ability. These categories are very coarse-grained and we did not get into the finer distinctions of different types of modality. Each of these categories are described in greater detail below and illustrated with examples.

When annotating the semantic tense, some of the events or states cannot be interpreted in relation to the document creation time and we have to annotate its relative tense. In such cases, we also link this text span to another that it depends on for its temporal interpretation. These links are all in the direction from the *dependent* text span to its *head* span. Such dependent text spans are typically tagged with one of the relative tense categories that include *Relative Present*, *Relative Past* or *Relative Future* when annotating *tense*.

As a practical matter, we also need to determine whether a text span needs to be annotated for all three dimensions in the first place. In (5a), for example, text spans that consist entirely of auxiliary verbs do not need to be annotated if they are followed by a main predicate and the tense and modality annotation will be associated with the main predicate. In (5b), the text span is a modal verb, and modal verbs only need to be annotated for tense, but not for event type and modality. In some cases, a verb as part of a Chinese verb compound is mapped to an adverbial particle or a preposition in English and such particles or prepositions need not to be annotated as the tense and modality is attached to the main verb in the compound. An example is given in (5c), and the main verb is italicized and the preposition is underlined.

- (5) a. Six months after the kidnapping , he still had n't *gotten the surgery* he needed to heal his burned flesh.
- b. You can make a contribution at : International Catholic Migration Commission Citibank USA 153 East 53rd Street , 16th floor New York, NY 10043 .
- c. Very few organizations are working on *getting* aid to Iraqi refugees.

2.1 Tense

We set up six categories for semantic tense, and these are *past*, *present*, *future*, *relative past*, *relative present*, and *relative future*. The relative tense

categories are inspired by the discussions of tense in (Comrie, 1985), and they are triggered by verbs in non-finite forms that have a clear dominating verb (“head verb” in dependency structure terms) that it depends on for its temporal interpretation. A link will also be annotated between the dependent verb and its head verb. The annotation of links will be discussed in greater detail in Section 2.4. Each of these six categories are defined and illustrated below. The relevant text spans are underlined.

Past The text span describes an event or state that happened in the past.

- (6) He started an engineering firm and worked with contractors such as ABB and Kellogg , Brown and Root;

Present The text span describes a present event or state. This includes a present state, an event that happens repeatedly in the present, a present on-going event, a completed event that has present relevance.

- (7) It is centered on the Hongshui River hydroelectric plant .

Future The text span describes an event that will happen in the future, or a future state.

- (8) Some people will prefer that option because it 's more convenient .

Relative Past The text span describes an event that happened in the past, or a past state relative to the event it depends on. In (9), “crossing” happened before the time “repeated” happened, i.e. “crossing” is relatively past to “repeated”.

- (9) After crossing a 30 - foot no man’s land we *repeated* the process at the second wall .

Relative Present The text span describes an event that happens in the present, or a present state relative to the event it depends on. In (10), “taking up” happens at the same time with “‘ve got”, i.e. “taking up” happens at the present relative to “‘ve got”.

- (10) I 've got two dead monitors taking up space in my office .

Relative Future The text span describes an event that happens in the future or a future state relative to the event it depends on. In (11a), “to strengthen” depends on “has invested” for its temporal interpretation, and “to create” depends on “to strengthen”. In (11b), “to be listed” depends on “approved” and the former logically occurs after the latter.

- (11) a. It *has invested* more than 130 billion yuan to strengthen the construction of infrastructures and basic industries so as to create a sound environment for expanding the opening up to the outside world . (link from “to strengthen” to “has invested”, from “create” to “has invested”, and from “for expanding” to “create”)
- b. Among them , 57 items *were approved* to be listed in the national , provincial and municipal Torch Plan and their quantity ratio is tops among the new , high level technology industry zones of the entire country .

Even in English, annotating tense can be challenging in at least two scenarios, and the first one being when there is a mismatch between the grammatical tense and the semantic tense. In (12), for example, “reaches” has a grammatical present tense, but it should be interpreted an event that occurs in the future. In this case, the grammatical tense can be deceiving and can be an impediment that prevents the annotator from making the correct decision. The other scenario is when the text span is a verb that takes on a non-finite form of a verb (13a) or other grammatical categories such as nouns (13b) and adjectives (13c). When this happens, tense is grammatically under-specified even in English, just like in Chinese. In this case, the temporal interpretation of the event depends on the larger context rather than the event denoting verb itself. While event denoting verbs in non-finite forms such as infinitives and participles can be annotated with a relative tense with a link to its dominating verb, event denoting nouns and adjectives often do not have one single dominating event that it can get its temporal interpretation from. Therefore, for nouns and adjectives as well as participle forms modifying nouns, we do not annotate relative tense. Instead, we assign an abso-

lute tense value that reflects the temporal interpretation of the event based on the context of the event. In (13b), for example, “acquittal” gets a past tense due to clues like “since”, and “conviction” gets a past interpretation due to the fact that it has a temporal modifier “1999”, which makes the temporal location of “conviction” explicit. In (13c), “rich” and “doting” are adjectives that modify nouns, and their temporal interpretation comes from the verb “appeared”. These adjectives are translated from predicative adjectives in relative clauses in Chinese, which can also be interpreted for tense based on the larger context.

- (12) To ensure that the money reaches the Iraqi program , write Iraq - icmc on your check.
- (13) a. This should be a motif familiar to anyone acquainted with the literature of mind control and ritual abuse survivors : the father and first controller , passing his child - victim up the social ladder of abuse in return for status , protection and reward .
- b. He ’ s moved on since his acquittal , like Gary Glitter did after his 1999 conviction , having departed last June for a Bahrain “ vacation ” from which he ’ s yet to emerge .
- c. For the people working at Bahrain ’ s malls , the person covered head to toe in a black veil , gloves and glasses appeared to be a rich , doting Saudi mother.

2.2 Eventuality Type

We define five eventuality types, and these are *habitual event*, *state*, *on-going event*, *completed event*, and *episodic event*. The eventuality type is set up as a way to help infer tense. Habitual events, on-going events, and states, for example, tend to occur in the present by default, while episodic events tend to occur in the past by default (Smith and Erbaugh, 2005). Given that there is no grammatical tense in Chinese, such a classification may prove to be an important source of information that helps predict tense. Each of the five types is described and illustrated below, and the relevant text spans are underlined:

Habitual Event The text span describes an event that happens repeatedly on a regular basis (14a, 14b). Habitual events are compatible with adverbial modifiers such as “often”, “usually”, “rarely”, “generally”, “seldom”, etc. Habitual events describe a pattern of actual events. General truths and statements also belong to this category (14c, 14d).

- (14) a. I used to drive to work but now I take the bus.
- b. At present , the Pu Kang Company , which produces the vaccine in this zone , has already formed a production scale of 5 million doses per year .
- c. Time flies.
- d. The moon travels around the earth.

State The text span describes an unchanging situation that will continue unless something happens to change it.

- (15) a. It is centered on the Hongshui River hydroelectric plant .
- b. but the demand far outstrips the money available to us , says Magy Mahrous , who oversees the project.
- c. According to investigation , each enterprise entering this zone has one or more new , high level technology projects or products .

On-going Event The text span describes an event that is on-going. The progressive aspect marker is generally a good indicator of this type of event.

- (16) a. At the school, where Bush was reading a story to a group of second-graders, the news came on TV that a second jet had hit the World Trade Center .
- b. God is testing us , he said .

Episodic Event The text span describes a situation that involves some sort of change or occurrence.

- (17) a. The National Weather Service reported that two other tornadoes touched down in the region - one in east Lindale, another in south-east Calhoun.

- b. Gross domestic product, the broadest measure of the nation’s economic growth, contracted at an annual rate of 0.1% from October to December, the Commerce Department said Wednesday.

Completed Event The text span describes a past event that has present relevance.

- (18) Within three to five years , Beihai has constructed the framework of a modernized city.

2.3 Modality

This dimension is used to distinguish events that actually happens from events that are intended, expected, possible, required, hypothetical. We define four modality categories – *actual event* is for events that actually happens, while non-actual event types include *intended event*, *hypothetical event*, and *modalized event*. These are described and illustrated below:

Actual Event The text span describes an event or state in the real world that actually happened, happens, is happening or will happen. This includes habitual events that happen repeatedly, or negated event that actually do not happen.

- (19) Beihai has already become a bright star arising from China ’s policy of opening up to the outside world .

Intended Event The text span describes an intended or expected event or state that does not necessarily happen or hold in the real world. This covers events that are intended, expected, planned, etc. An intended event typically follows a main verb and denotes the purpose or intention of the main verb. Those text spans are typically verbs in non-finite forms and are linked to a main verb that occurs before or after it. Text spans following modal verbs are excluded from this category, and are put into *Modalized event* category.

- (20) a. It has also drafted three documents for attracting foreign capital , strengthening horizontal economic integration and allowing more authority for foreign operations .

- b. Among them , 57 items were approved to be listed in the national , provincial and municipal Torch Plan and their quantity ratio is tops among the new , high level technology industry zones of the entire country

Hypothetical Event The text span describes an event or state that is in a conditional (e.g., if, when) clause or takes place conditional on something else, and does not necessarily happen in reality.

- (21) Would the experiment have been as successful if they had not spent the money ?

Modalized Event The text span follows a modal verb, and describes a possible or necessary event or state, or an ability.

- (22) a. The recent confrontation could ignite regional convulsions as Turkey is sucked into Syria, leading to belated actions from the international community.
- b. That now will not happen, but it is possible that he could be summoned by Congress to testify later.

Our annotation scheme for modality as it currently stands is still very coarse-grained. For example, we do not distinguish the different types of modalities traditionally introduced by modal verbs, such as epistemic and deontic modals. The classification is also shallow in that our intent is to simply identify syntactic constructions that have a modal interpretation. For instance, the sentences in (23) definitely expresses uncertainty, but because they are not associated with one of the above syntactic constructions, they are still considered to be “actual”, which is the default category for modality.

- (23) a. It will probably rain tomorrow.
- b. It will possibly rain tomorrow.

2.4 Links

The annotation of links is triggered by events expressed by verbs in non-finite forms that have a clear dominating (head) verb. The annotation of links is closely tied to the annotation of relative tense. When an event is annotated with a relative tense category,

a link is annotated so that the relative tense of the dependent event can be interpreted in relation to the temporal location of the dominating verb. The link is always in the direction from a dependent text span to a head text span.

- (24) a. To further expand the opening up to the outside and promote outwardly economic development , Guangxi has come up with a series of policies to make use of foreign investments .
- b. This development zone is located in the downtown area of Hangzhou , a famous Chinese scenic sightseeing city , and is a national level new , high level technology industry development zone approved for construction by the State Council in 1991 .

In (24a), “expand” and “promote” are both linked to “has come up with”, with relative future tense assigned to the first two spans. In (24b), “for construction” is annotated with a relative future tense and linked to “approved”.

3 Annotation Experiments

As of this writing, we have completed the first round of our annotation experiments. We selected 50 sentences, which consist of 944 words, from the Parallel Aligned TreeBank (Li et al., 2012) from the LDC. There are 167 text spans that are marked up as event anchors. The annotation experiment involves three annotators. Each sentence is annotated three times and we computed their pairwise agreement statistics. The results of the average agreement scores are presented in Table 1:

	Tense	Event Type	Modality
Agreements	78.6% (131/167)	73.5% (123/167)	81.4% (136/167)
Kappa scores	0.71	0.65	0.70

Table 1: Inter-annotator agreement

It is premature to draw any firm conclusions about the effectiveness of this distant annotation approach with this first round of annotation, but it is worth noting that the inter-annotator agreement statistics

are already comparable with and even better than that reported in (Xue et al., 2008) where tense is directly annotated on Chinese text using annotators that have undergone significant training and there were a smaller number of tense categories (four). We believe this shows the initial promise of this approach. From the statistics in Table 1, we can also see that our annotators show better agreement for Modality and Tense than Eventuality Type. We analyzed our annotation results through confusion matrices which show that the most difficult distinction for Tense is between *Present* and *Future*, and the confusion happens mostly on modal verbs, modalized events and hypothetical events, where it is difficult to distinguish these two tenses. The most challenging distinction for Eventuality is between *Habituals* and *State*, indicating that the distinction between those tags is still vague and not clearly defined.

4 Related Work

In a series of TempEval evaluations (Verhagen et al., 2007; Verhagen et al., 2010) that are aimed at detecting time expressions, events and the relations among them, (abstract) tense determination is formulated as a task of determining the relation between an event and the document creation time. TempEval uses a fairly coarse-grained set of values (Before, Before or overlap, Overlap, After, After or Overlap) to represent abstract “tense”. The “tense” annotation task in the TempEval evaluations targets the main event of a sentence, while we are attempting to annotate the semantic tense for all eventualities, including events and states in non-finite verb forms or even in nominal forms. We also define an *event type* classification that is intended to help infer tense, mindful of the fact that there will not be morpho-syntactic cues on the Chinese side that can help make such determination. We also attempt to set up a fairly coarse-grained classification system for modality of events. Tense, aspect, and modality are also annotated in the TimeBank (Pustejovsky et al., 2005) as attributes of events, but TimeBank generally annotates the grammatical tense of English verbs. For example, the tense of an event that takes the form of a non-finite verb or a noun will get the value of “None” even though the semantic tense for some non-finite verbs

can be determined, as shown in (4). We target the semantic tense instead of the grammatical tense because we think they are “transportable” across languages. Our ultimate goal is to infer the semantic tense on the Chinese side, not just on the English side.

A recent attempt to annotate and disambiguate the semantic tense for English is by Reichart and Rappoport (2010), who introduced a more general Tense Sense Disambiguation (TSD) task that provides a fine-grained sense taxonomy for tense. They view tense as having three different levels: Concrete (surface) Syntactic Forms (CSF, e.g., am/is/are V-ing), Abstract Syntactic Forms (ASF, e.g., present progressive), and a taxonomy of 103 different underlying senses. For example, Reichart and Rappoport define 11 underlying senses for the “simple present” ASF which include “things that are always true”, “general and repeated actions and habits”, “plans, expectations and hopes”. These fall under the scope of modality in theoretical linguistics research (Kratzer, 1981; Carlson and Pelletier, 1995; Guéron and Lecarme, 2008). Reichart and Rappoport essentially use modality as a semantic dimension to disambiguate the different “senses” of abstract tense forms. Their goal is to predict the underlying senses given the surface CSFs. We target similar distinctions in our annotation, but instead of treating these distinctions as unstructured fine-grained senses, we classify events along three different dimensions that in conjunction can make similar distinctions in a more structured manner.

On the Chinese side, there have been several past attempts to infer “tense” for Chinese automatically using statistical models and modest success has been reported. There are two general approaches to “tense” inference for Chinese. The first approach has been to manually annotate tense on Chinese verbs (Ye et al., 2006; Ye, 2007; Xue et al., 2008; Xue, 2008) and use the annotated data to train statistical models to predict tense in previously unseen text. (Xue, 2008) has shown that even though there are no morpho-syntactic clues for tense in Chinese, contextual information can be exploited to infer “tense”. Such contextual information includes explicit clues such as time expressions and aspect markers as well as implicit information such as verb types: bounded events (e.g., “explode”) tend to oc-

cur in the past while unbounded events tend to occur in the present (e.g., “believe, know, like”), a generalization first articulated in (Smith and Erbaugh, 2005). Maintaining consistency among annotators when annotating a phenomenon with a total lack of explicit surface cues, however, proved to be a very challenging task. (Xue et al., 2008) reported an inter-annotator agreement of 75% despite of using a fairly coarse-grained tagset (Xue et al., 2008), a result that is comparable to that of our first round of annotation. The second approach is cross-lingual projection and this is the tack that Liu et al (2011) took. They mapped grammatical tense in English onto Chinese via word-aligned parallel text. The issue with mapping the surface grammatical tense is that, as we discussed above, the syntactic forms of tense are ambiguous with regard to their underlying semantics. As a result, when they are projected onto a different language, the same context will point to different grammatical tense categories. This will confuse the statistical machine learning models and hamper their performance and hence limit the utility of the resulting automatic systems.

We believe that our distant annotation approach combines the best of both worlds. Unlike manual annotation on just the target language (Chinese) side, we benefit from the presence of the morpho-syntactic cues in the source language (English). Effectively, our distant annotation approach allows us to annotate the underlying semantics of tense in the easier source language and map it to the more difficult target language. At the same time, unlike direct projection of surface forms of tense, our distant annotation approach maps the surface forms to an underlying semantic representation that is free from language-specific idiosyncracies at the morpho-syntactic level.

5 Conclusions and Future Work

We describe a distant annotation approach for annotating the tense, event type and modality of events in Chinese text by annotating their English counterpart via a word-aligned parallel corpus. Preliminary results indicate that this approach shows promise as an effective alternative to annotating the Chinese text directly, a challenging task since Chinese does not have the morpho-syntactic cues that constrain anno-

tation choices. We are currently performing additional annotation experiments while refining our annotation guidelines. The ultimate goal is to generate consistently annotated data on the Chinese side that can be used to train statistical models to automatically predict tense, event type and modality of Chinese events. We believe such a tool would benefit a wide variety of natural language applications that include Machine Translation and Information Extraction.

Acknowledgment

This work is supported by the IIS Division of National Science Foundation via Grant No. 0910532 entitled “Richer Representations for Machine Translation” . All views expressed in this paper are those of the authors and do not necessarily represent the view of the National Science Foundation.

References

- Gregory N. Carlson and Francis Jeffrey Pelletier. 1995. *The Generic Book*. Chicago: University of Chicago Press.
- Bernard Comrie. 1985. *Tense*. Cambridge University Press.
- Michelle Cutrer. 1994. *Time and Tense in Narratives and Everyday Language*. Ph.D. thesis, University of California in San Diego.
- John Dinsmore. 1991. *Partitioned Representations*. Dordrecht, Netherlands: Kluwer.
- Gilles Fauconnier. 2007. Mental Spaces. In Dirk Geeraerts and Hubert Guyckens, editors, *The Oxford Handbook of Cognitive Linguistics*.
- Jacqueline Guéron and Jacqueline Lecarme. 2008. *Time and Modality*. Springer.
- Angelika Kratzer. 1981. The notional category of Modality. In H.J. Eikmeyer and H. Rieser, editors, *Words, Worlds and Contexts*. Berlin: de Gruyter.
- Xuansong Li, Stephanie Strassel, Stephen Grimes, Safa Ismael, Mohamed Maamouri, Ann Bies, and Nianwen Xue. 2012. Parallel Aligned Treebanks at LDC: New Challenges Interfacing Existing Infrastructures. In *Proceedings of LREC-2012*, Istanbul, Turkey.
- Feifan Liu, Fei Liu, and Yang Liu. 2011. Learning from chinese-english parallel data for chinese tense prediction. In *Proceedings of the 5th International Conference on Natural Language Processing*, pages 1116–1124, November.

- James Pustejovsky, Bob Ingria, Roser Sauri, Jose Castano, Jessica Littman, Rob Gaizauskas, Andrea Setzer, G. Katz, and I. Mani. 2005. The specification language TimeML. In I. Mani, J. Pustejovsky, and R. Gaizauskas, editors, *The Language of Time: a Reader*. Oxford University Press.
- Roi Reichart and Ari Rappoport. 2010. Tense sense disambiguation: A new syntactic polysemy task. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 325–334, Cambridge, MA, October. Association for Computational Linguistics.
- Hans Reichenbach. 1947. *Elements of Symbolic Logic*. The MacMillan Company, New York.
- Carlota S. Smith and Mary Erbaugh. 2005. Temporal interpretation in Mandarin Chinese. *Linguistics*, 43(4):713–756.
- Carlota Smith. 2001. Discourse modes: Aspectual and tense interpretation. In J. Busquets, F. Lambert, and A. Le Draoulec, editors, *Cahiers de Grammaire 26: Sémantique et Discours*. ERSS et Université de Toulouse-le Mirail, Toulouse.
- Marc Verhagen, Robert Gaizauskas, Frank Schilder, Mark Hepple, Graham Katz, and James Pustejovsky. 2007. Semeval-2007 task 15: Tempeval temporal relation identification. In *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, pages 75–80, Prague, Czech Republic, June. Association for Computational Linguistics.
- Marc Verhagen, Roser Sauri, Tommaso Caselli, and James Pustejovsky. 2010. Semeval-2010 task 13: Tempeval-2. In *Proceedings of the 5th International Workshop on Semantic Evaluation*, pages 57–62, Uppsala, Sweden, July. Association for Computational Linguistics.
- Nianwen Xue, Zhong Hua, and Kai-Yun Chen. 2008. Annotating tense in a tenseless language. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation*, Marrakech, Morocco.
- Nianwen Xue. 2008. Automatic Inference of the Temporal Location of Situations in Chinese Text. In *EMNLP-2008*, Honolulu, Hawaii.
- Yang Ye, Victoria Li Fossum, and Steven Abney. 2006. Latent features in automatic tense translation between Chinese and English. In *The Proceedings of the 5th SIGHAN Workshop on Chinese Language Processing*, Sydney, Australia.
- Yang Ye. 2007. *Automatic Tense and Aspect Translation between Chinese and English*. Ph.D. thesis, University of Michigan.