

If it were *then*, then when was it? Establishing the anaphoric role of *then*

Georgiana Puşcaşu*, Ruslan Mitkov

University of Wolverhampton, UK

{georgie, r.mitkov}@wlv.ac.uk

Abstract

The adverb "then" is among the most frequent English temporal adverbs, being also capable of filling a variety of semantic roles. The identification of anaphoric usages of "then" is important for temporal expression resolution, while the temporal relationship usage is important for event ordering. Given that previous work has not tackled the identification and temporal resolution of anaphoric "then", this paper presents a machine learning approach for setting apart anaphoric usages and a rule-based normaliser that resolves it with respect to an antecedent. The performance of the two modules is evaluated. The present paper also describes the construction of an annotated corpus and the subsequent derivation of training data required by the machine learning module.

1. Introduction

The text understanding mission relies, among other key features, on the ability to interpret temporal information. A correct temporal interpretation leads to improvement in the performance of many NLP applications. Temporal information is conveyed in many ways in natural language: tense, aspect and lexical items that carry temporal information, such as temporal expressions (TEs).

TEs play a very important role in the temporal interpretation of text. They not only convey temporal information on their own, but also serve as anchors for locating events referred to in text. But certain TEs are ambiguous, in the sense that they either have different temporal interpretations (e.g. "today" can be used to denote the day of the utterance, but also with the generic sense "nowadays"), or they can express more semantic roles (e.g. "then" can play the role of a linking adverbial, but also realize the semantic role of time). The process of anchoring TEs on a timeline is called temporal resolution or normalisation.

In this paper we will focus on the disambiguation and temporal resolution of the adverb "then". More specifically, this paper will report on an empirical investigation of all possible usages of "then", as well as on the design and evaluation of an algorithm aiming to set apart the co-temporal anaphoric usage of "then" and resolve it relative to an antecedent. The individual study of "then" in the context of temporal resolution can be resembled to the individual study of "it" in the anaphora resolution process.

The paper is structured as follows: section 2 motivates our intentions to recognise the anaphoric usage of "then" and surveys related work, section 3 explores the phenomenon with respect to its grammatical characteristics and delimits five profiles to be employed in the classification of "then". In section 4, the development of a novel corpus for use in training and evaluation is described. A machine learning approach to recognition of anaphoric "then" is proposed and evaluated in section 5. Section 6 directs its attention towards the normalisation of anaphoric "then" by identifying the antecedent it refers to. Finally, in section 7 conclusions are drawn and future research considered.

2. Motivation and related work

In recent years, the task of temporal expression recognition and normalisation has received increased attention: (Mani and Wilson, 2000), (Setzer and Gaizauskas, 2000), (Filatova and Hovy, 2001), (Katz and Arosio, 2001), (Schilder and Habel, 2001), (Puscasu, 2004). The importance of the proper treatment of TEs is reflected by the relatively large number of NLP evaluation efforts centered on their identification and normalisation, such as the MUC 6 and 7 Named Entity Recognition tasks (MUC-6, 1995), the ACE-2004 Event Recognition task (ACE, 2004), the Temporal Expression Recognition and Normalisation task (TERN, 2004).

The identification and normalisation of TEs is a prerequisite for capturing the temporal dimension of a given text. The process of resolving under-specified temporal expressions requires finding the anchor (an already resolved TE) that the time denoted by the expression is relative to. Our recent research (Puscasu, 2004) has focused on the development of a temporal tagger capable of identifying both self-contained TEs, which get tagged with their value, and indexical / under-specified TEs, which, depending on their semantics, receive a value computed by a temporal function having as argument the time they are relative to.

A relevant source of errors we discovered in the normalisation process is providing a temporal anchor to a multivalent TE that appears in text without making any reference to a specific point in time. This fact was confirmed by Mani and Wilson (2000), who trained a classifier to distinguish specific usage of "today" (meaning the day of the utterance) from its generic usage meaning "nowadays", illustrating one challenge posed by these errors, that is distinguishing between specific use and generic use of the same TE (e.g. "today", "now"). Another challenge is the temporal adverbial "then", which, as corpus study reveals (Biber et al., 2000), is the second most frequent temporal adverbial appearing in English texts, the first being "now".

"Then" is an adverb of great communicative strength, easily expressing one or another semantic category (or more than one simultaneously). The adverb "then" can either refer to a time given in the context (synonym with "at that time" anaphoric usage), or, quite commonly, mark the next event in a sequence, denote a result/inference or mark enumerations.

* Now visiting researcher at the University of Alicante, Spain

tions, as well as antithesis. Only the first usage of "then" should receive a temporal value, but the second use is also important for the task of temporally ordering events. The accurate recognition of a particular usage of "then" thus contributes to all fields in which temporal information is a concern, whether it be event-based information organization, text summarisation or question answering.

Whereas in recent years research in temporal information processing has been enjoying growing attention and has produced some encouraging results, the computational treatment of the anaphoric use of "then" has not been properly addressed yet. Hitherto, a theoretical study of "then" worthwhile mentioning was performed by Glasbey (1993) who designed a grammar based on semantic theoretic discourse representation theory in order to generate a particular reading for sentence-final "then". From another angle, our paper attempts to identify the anaphoric use of "then" without recourse to semantic information, and to implement a procedure which will instantiate its temporal value.

3. A grammatical overview of "then"

The adverb "then" receives coverage in most serious surveys of English grammar, including (Quirk et al., 1985) and (Biber et al., 2000).

From a syntactic point of view, Biber et al. (2000) distinguishes between "then" as complement of prepositions (*since then, before then*) and "then" as a clause element (adverbial). As adverbial, "then" either adds information about the action described in the clause (functioning as a **circumstance adverbial**) or connects stretches of text (**linking adverbial**). Corpus findings (Biber et al., 2000) certify the fact that "then" is the most frequent linking adverbial, this making our task more difficult, as anaphoric "then" functions syntactically as a circumstance adverbial.

From a semantic point of view, "then" can either express the semantic category of time, indicating time-position or temporal relationship, or, as linking adverbial, it can have a considerable range of meanings, such as inferential, summative, enumerative or antithetic. Quirk et al. (1985) confirms the considerable range of meanings "then" can have, but also infers that the central core of meaning remains temporal. "Then" can often be paraphrased by "at that time" or "after that time". The aim of our present study is to set apart "then" standing for "at that time", as in:

For centuries archaeologists have argued over descriptions of how Archimedes used concentrated solar energy to destroy the Roman fleet in 212BC. Historians have said nobody then knew enough about optics and mirrors. (Mitkov, 2002)

3.1. Guidelines and profiles for classification

By understanding and exploring the full range of phenomena, we can better characterize anaphoric usages of "then" and devise a more accurate algorithm for identifying them. Surveying the uses of "then" in English, both as described in the literature and by an examination of the corpus described below in section 4, allows the identification of five different profiles for this adverb. The five profiles will correspond to the classes furtherly used in the classification process, therefore are given symbolic names. They are spelt out with examples below:

1. ANAPHORIC

* "Then" functions as complement of prepositions (such as "by", "since", "from", "until", "beyond", "around")
New Delhi exploded a nuclear device in 1974, but has not undertaken any nuclear tests since then.

* It has the syntactic function of circumstance adverbial and denotes a point or a period of time specified before (synonym with "at that time").
They lived in London for the first few years of their marriage and were then very happy.

* "Then" is used to correlate with a conjunction that introduces a preceding subordinate temporal clause.
When war actually came, then the country started to panic.

2. TIME_REL

* "Then" functions as a circumstance adverbial of time indicating time relationship or denoting temporal sequence (synonym with "afterwards").
The state has to hold 51 percent of Lietuvos Nafta for three years but can then bring its share down to 34 percent.

3. INFERENCE

The adverb "then" functions syntactically as a linking adverbial, but semantically it can have one of the following meanings:

- **result/inference**, as in:

* *if*-clauses, followed by a correlating inferential "then"
"One of the great lessons of history is that if America is prepared to fight many wars and greater wars and any wars that come, then we will fight fewer wars and lesser wars and perhaps no wars at all", said Dole.

* correlative usages (where "then" correlates with subordinators other than "if" that introduce a preceding non-temporal clause) or any other usage of "then" as synonym with "therefore".
Because Jennifer foresaw this, she then had the time to change her plans.

* comments of an inferential nature
She is not at the cinema. Where did she go then?

- **summation**

* "then" is used with the summarizing sense also captured by "altogether", "therefore" or "thus"
- *He lost his watch, his car broke down, he got a letter of complaint from a customer...*
- *Then we can say he had a bad day!*

4. ENUMERATIVE

* The adverb "then" functions syntactically as a linking adverbial, but semantically it is used to structure items in a list.
First of all it's a cold day. Then there are clouds. Thirdly there is fog. These are some of the reasons I prefer not to travel today.

* Aside from structuring elements in a list, it is used as reinforcement, giving higher weight to an item in a list.
He has the opportunity, the motivation, and then the courage to do it.

	ANAPHORIC	TIME_REL	INFERENTIAL	ENUMERATIVE	ANTITHETIC	ERROR	SUM
ANAPHORIC	307	14	7	0	0	0	328
TIME_REL	9	542	13	12	2	0	578
INFERENTIAL	4	17	193	0	1	0	215
ENUMERATIVE	0	15	0	17	0	0	32
ANTITHETIC	2	4	3	0	0	0	9
ERROR	0	0	0	0	0	11	11
SUM	322	592	216	29	3	11	1173

Table 1: Interannotator agreement for six usages of "then"

	ANAPHORIC	NON-ANAPHORIC	SUM
ANAPHORIC	307	21	328
NON-ANAPHORIC	15	830	845
SUM	322	851	1173

Table 2: Interannotator agreement for two usages of "then"

5. ANTITHETIC

* "Then" is used as linking adverbial to present something in contrastive relation to something preceding it (synonym with "on the contrary", "on the other hand").

You promise to help me, then you let me down!

The five categories previously distinguished will form the basis of our experiments. Even though the purpose of this paper is to delimit only the ANAPHORIC set of uses, the present investigation also treats the other four categories trying to design a methodology to assign a category to each appearance of "then" in text.

4. Corpus Annotation

The corpus study will be performed on newspaper articles included in the Reuters Corpus (2000). A corpus was constructed by random selection of 1,000 articles, so that the word "then" appears at least once within each document. In total, the annotated data contained 410,391 words, with 1,173 occurrences of "then". The corpus has been annotated by two different annotators, in order to measure the interannotator agreement, thus gaining an insight into the complexity of the problem and the validity of the designed categories (see subsection 3.1). To facilitate the markup of the usage type of "then", only paragraphs containing the word together with one preceding paragraph (extracted to provide context) have been presented to the annotators. Each human annotator has been asked for a decision about the class "then" belongs to. The annotators had to decide among six classes: ANAPHORIC, TIME_REL, INFERENTIAL, ENUMERATIVE, ANTITHETIC and ERROR. The class ERROR has been introduced as cases have been observed during annotation where "then" was incorrectly used instead of "than" due to typing errors. The class ERROR is thus assigned to an appearance of "then" as in the following context:

There was about 1,400 to 1,600 people at the protest today - much less than last week.

4.1. Interannotator agreement

Reliable annotated data is necessary for a wide variety of natural language processing tasks. While the objective correctness of human annotations cannot be computationally judged, the degree to which the annotators agree in their labelling of a corpus can be statistically determined using the kappa measure (Siegel and Castellan, 1988). Commonly, the kappa statistic is used to measure interannotator agreement. It determines how strongly two annotators

agree by comparing the probability of the two agreeing by chance with the observed agreement. If the observed agreement is significantly greater than that expected by chance, then it is safe to say that the two annotators agree in their judgements.

Mathematically,

$$K = \frac{P(A) - P(E)}{1 - P(E)} \quad (1)$$

where P(A) is the proportion of times that the annotators agree and P(E) is the proportion of times that we would expect them to agree by chance. Krippendorff (1980) discusses what makes an acceptable level of agreement, saying that content analysis researchers generally think of $K > 0.8$ as good reliability, with $0.67 < K < 0.8$ allowing tentative conclusions to be drawn.

In our case, the agreement matrix resulted by analyzing the two annotations is presented in Table 1.

The proportion of agreement observed in corpus annotation is 91.22% and the proportion of agreement expected due to chance is 36.00%. Applying formula 1, we obtain a kappa agreement between the two annotators of 86.28%.

As one can easily deduct from 1, the annotators have never agreed on antithetic usages of "then", leading to the conclusion that either guidelines for the class ANTITHETIC are not well defined or the antithetic value always overlaps with other semantic values, being difficult to set apart. The capacity of "then" to express more semantic categories simultaneously accounts for many differences between the opinions of the two annotators, as exemplified below:

* TIME_REL vs. INFERENTIAL

The way the polymer and the dye mixtures reacted to the vapours changed the light signal, and the researchers could then measure and categorise the change caused by different smells.

* TIME_REL vs. ENUMERATIVE

"First they could put on the quota, and then there would be the question of decreasing the size," he said.

* INFERENTIAL vs. ANAPHORIC

Parolin tried to beat the beast with a stick and the cougar then turned on her; the Royal Canadian Mounted Police said.

* INFERENTIAL vs. ANTITHETIC

Russian banks are generally selling, with foreigners buying immediately, but then prices are going down without trading...

Since our main target is identifying only the anaphoric usage of "then", we have also measured the interannotator agreement when distinguishing only between two types of usages: anaphoric and non-anaphoric. We have thus included all but the anaphoric usage (time relationship, inferential, enumerative, antithetic and erroneous) within the same class called NON-ANAPHORIC. Table 2 illustrates the agreement matrix obtained by setting apart the anaphoric from the non-anaphoric usage. As one can easily deduct from the table, one annotator has encountered 322 anaphoric "then"s and the other one 328, out of which only 307 were commonly agreed on.

Making the distinction only between two classes, the proportion of agreement observed in corpus annotation is 96.93% and the proportion of agreement expected due to chance is 59.94%. The kappa agreement between the two annotators is therefore in this case 92.33%.

5. Machine learning applied to the classification of "then"

Machine learning has been successfully employed in solving many NLP tasks. Evans (2000) used a methodology similar to that used in this work for a similar task, the identification of non-nominal "it". As in the present paper, a human annotated corpus served to supervise the machine learning method. In that specific task, the author has performed a comparative evaluation of a rule-based and a machine learning approach, obtaining with both approaches similar results.

The machine learning method we selected to apply to the problem discussed in this paper is presented in subsection 5.1, then the set of features used for classification follows in subsection 5.2. Subsection 5.3 comprises the experimental setup and a discussion of results.

5.1. Memory based learning

Memory based learning (MBL) is a supervised inductive learning algorithm for solving classification tasks. It has proven to be successful in a large number of tasks in the domain of natural language processing. MBL is based on the idea that intelligent behaviour can be obtained by analogical reasoning, rather than by the application of abstract mental rules as in rule induction and rule-based processing. In particular, it is founded on the hypothesis that the extrapolation of behaviour from stored representations of earlier experience to new situations, based on the similarity of the old and the new situation, is of key importance. MBL algorithms take a set of examples (fixed-length patterns of feature-values and their associated classes) as input, and produce a classifier which can classify new, previously unseen, input patterns.

The MBL algorithm we used for learning, and then classifying, is k-nearest neighbours. The training set contains instances characterised by a succession of feature values and the class associated to that instance. At the training stage, the training instances are treated as points in a multi-dimensional feature space and stored as such in an instance base in memory. When the trained classifier is confronted with a test instance characterised by a set of feature values, a distance metric is used to compare the position of

the test instance with respect to all training instances in the feature-defined multi-dimensional space. The closest k-nearest neighbours are selected and the test instance is then classified based on the most frequent classification of the k selected neighbours.

For the purposes of the work described in this paper we have used the implementation of k-nearest neighbours included in the software package called TiMBL (Daelemans et al., 2004). Each training and test instance has been characterized by 20 features described below.

5.2. Feature description

In order to classify instances of "then" into one of the six classes established above, a set of 20 features was defined. The features were defined so that their values can be automatically extracted from any text analysed with Conexor's FDG Parser (Tapanainen and Jaervinen, 1997). This parser returns information on a word's part of speech, morphological lemma and its functional dependencies on surrounding words.

The current "then" classifier employs a set of 20 features detailed below:

1. **Position** This attribute defines the position of "then" with respect to the closest clause subject and predicate. It can have the following values:
 - * *Initial* - "then" appears before the closest subject and predicate;
 - * *Initial medial* - "then" follows the subject, but precedes the predicate;
 - * *Medial* - it is situated between two auxiliaries of the same predicate;
 - * *End medial* - "then" follows the auxiliary or the infinitive particle, but comes before the main verb
 - * *End* - it follows both the subject and the predicate.
2. **POS-2** denotes the part of speech of the token preceding the token coming immediately before "then" (punctuation marks are considered).
3. **POS-1** corresponds to the part of speech of the token preceding "then".
4. **POS** is the part of speech associated to "then".
5. **POS+1** corresponds to the part of speech of the token following "then".
6. **LinkedTo** represents the part of speech of the word "then" is linked to.
7. **PossiblyInsideNP** indicates whether "then" is possibly included within a noun phrase.
8. **TensePrecedingVP** corresponds to the tense of the predicate preceding "then".
9. **TenseFollowingVP** corresponds to the tense of the predicate coming after "then".
10. **DistancePrecedingVP** is the distance in tokens between "then" and the previous verb phrase (VP).
11. **DistanceFollowingVP** is the distance in tokens between "then" and the following VP.
12. **PrecededByAnaphoricInd** indicates whether or not "then" is preceded by prepositions that indicate a position in time, such as "from", "since", "before", "by", "until", "beyond", or by other indicators of anaphoric usage (e.g. "between now and").

13. **PrecededByAdverbs** indicates if "then" is preceded by adverbs such as "even", "just", "only".
14. **PresenceIfThen** encodes the presence of "if", as well as the presence of another "then" in the same sentence, together with their relative position with respect to the occurrence of "then" under investigation.
15. **PresenceFirst** indicates whether "then" is preceded by indicators of enumerative usage, such as "first" or "firstly".
16. **PresenceTemporalSub** indicates if "then" is preceded by a subordinating conjunction that introduces temporal clauses.
17. **PresenceBecause** shows if "then" is preceded by "because" within the same sentence.
18. **PrecededCC** corresponds to the presence of a coordinating conjunction ("and", "but") immediately before "then".
19. **TypeSentence** encodes the type of the sentence "then" appears in: declarative, interrogative or exclamative.
20. **PredictedFunction** indicates the function of "then" predicted by the FDG parser.

5.3. Experimental setup and evaluation

The machine learning method we used for the classification of "then" requires training data. Here, the training data is derived from the annotated corpus described in section 4. It contains, for each appearance of "then", an ordered list of 21 values, out of which the first 20 elements are the corresponding values for the features described above and the 21st item represents the class assigned in the annotation process.

Due to the fact that the 1173 occurrences of "then" were labelled by two human annotators, for experimental purposes two training sets were created: the first is the result of the annotation of one randomly chosen human annotator (we will refer to this training set as ORIGINAL) and the second contains only those cases when both annotators agreed upon the class attached to "then" (we will name this set AGREED). The ORIGINAL training set therefore contains 1173 training instances, while the AGREED set comprises 1070 training instances.

As previously mentioned, the implementation of k-nearest neighbours included in TiMBL was used for experiments. Features were weighted by gain ratio, and overlap was the distance metric between values of the same feature. Different numbers of nearest neighbours were experimented with, but the best performance was achieved when k=5. The evaluation was performed with the *leave-one-out* approach, a reliable way of testing the real error of a classifier (Weiss and Kulikowski, 1991). The underlying idea is that every instance in turn is selected once as a test item, and then the classifier is trained on all remaining instances.

Table 3 presents, for the two training sets previously delimited (ORIGINAL and AGREED), the accuracy of three classifiers corresponding to three different values assigned to k (1, 3, respectively 5).

As our target is setting apart only the anaphoric usage of "then", both at the training and testing stages we made the same distinction as when we measured interannotator agreement. We have thus considered only the ANAPHORIC and NON-ANAPHORIC categories, the classes TIME-REL, INFERENTIAL, ENUMERATIVE,

ORIGINAL		
k = 1	k = 3	k = 5
81.15%	83.97%	84.05%
AGREED		
k = 1	k = 3	k = 5
84.76%	86.82%	87.75%

Table 3: Evaluation results for six usages of "then"

ANTITHETIC and ERROR being all covered under the generic NON-ANAPHORIC designation. As in the above described experiments, two training sets were considered, but this time the assigned labels were only two. These two training sets will be furtherly referred to as ORIGINAL_TWO and AGREED_TWO. Table 4 illustrates for both training sets the accuracy of three classifiers corresponding to three different values assigned to k (1, 3, respectively 5).

ORIGINAL_TWO		
k = 1	k = 3	k = 5
90.11%	90.62%	91.04%
AGREED_TWO		
k = 1	k = 3	k = 5
90.18%	90.56%	91.58%

Table 4: Evaluation results for ANAPHORIC vs. NON-ANAPHORIC "then"

6. Temporal resolution of anaphoric "then"

So far, results have shown us that anaphoric "then" can be detected with an accuracy of 91%. After detecting the co-temporal anaphoric usage of "then", our next aim was normalising it, that is finding the temporal expression that "then" corefers or substitutes in order to avoid a repetition. For this purpose, the temporal tagger presented in (Puscasu, 2004) was employed to identify and normalise the temporal expressions that, unlike "then", can always be anchored on a timeline.

An investigation of all appearances of anaphoric "then", with respect to its temporal value in text, has revealed the following:

- * out of 322 instances of anaphoric "then", 271 (84.16%) were found to substitute a temporal expression ([1]), 14 (4.34%) to stand proxy for temporal *when*-clauses ([2]), and 37 (11.49%) were referring to the time an event occurred ([3]).

[1] *In 1991, Taiwan lifted a ban on the hiring of foreign workers to meet what was then a severe domestic labour shortage...*

[2] *And when the World Court says we should starve all Iranians, then we will comply.*

[3] *The government must publish a new order with the ministries rates, and until then the tax service and customs must charge preliminary, calculated rates...*

was fully specified in 122 cases (45.02%) and under-specified in 149 cases (54.98%).

- * the most important aspect for the normalisation of anaphoric "then" is that, when it was substituting a temporal expression, that TE was the last one in text before "then" in 241 cases (88.93%), the one before the last in 20 cases (7.38%) and the one following "then" within the same sentence in 10 cases (3.69%). It is worthwhile mentioning that, out of the 20 cases "then" was referring to the second preceding TE, in 5 cases the temporal expression preceding "then" was "now", and in 7 cases it was denoting a duration.

Taking into consideration the results of this investigation, we have developed a simple normaliser for anaphoric usages of "then". The search for an anchor-TE was first conducted in the same sentence in the text preceding "then", then in the previous sentence, and, if no TE was found, in the part of the sentence following "then". If still no TE was found, the text anterior to the preceding sentence was considered. This normaliser obviously does not account for the cases "then" plays the role of a place-holder for a temporal clause or for the occurrence time of an event, but still encounters the correct anchor for 78.26% of anaphoric "then"s.

7. Conclusions

This paper discusses the disambiguation and temporal resolution of the adverb "then". An empirical investigation has shown a variety of uses "then" can have in English texts. Five main classes of usages were defined, with the paper focusing on the anaphoric one. A corpus was annotated by two annotators and the interannotator agreement was computed. An analysis of the agreement between annotators has shown that "then" is not easily analysed not even by humans.

Several classifiers were then trained on the resulted training data. The best performing classifier was able to set apart anaphoric use of "then" with an accuracy of 91.58%. Improving the performance of the classifier relies on increasing the size of the training data, but also on the formulation of more effective features.

Having detected each co-temporal anaphoric occurrence of "then", the next step was resolving it relative to an antecedent. A normaliser was developed according to a few simple rules established as a result of corpus investigation. This normaliser encounters the correct antecedent for 78.26% of anaphoric "then"s. Future research is envisaged to enhance the performance of the normaliser to be able to recognise "then" as place-holder for temporal clauses or for times associated to events, as soon as the state-of-the-art in the area of event identification makes it possible.

8. Acknowledgements

The authors would like to thank Verginica Barbu Mititelu and Duygu Can for help with corpus annotation. In addition, we would like to thank Constantin Orasan for his comments and support.

9. References

- ACE. 2004. <http://www.nist.gov/speech/tests/ace/ace04/index.htm>.
- D. Biber, S. Johansson, G. Leech, S. Conrad, and E. Finegan. 2000. *Longman Grammar of Spoken and Written English*. Longman.
- W. Daelemans, J. Zavrel, K. Slood, and A. Bosch. 2004. TiMBL Tilburg Memory Based Learner, version 5 1, Reference Guide. Ilk technical report 04-02.
- R. Evans. 2000. A Comparison of Rule-Based and Machine Learning Methods for Identifying Non-Nominal It. In *Proceedings of NLP 2000*.
- E. Filatova and E. Hovy. 2001. Assigning Time-Stamps to Event-Clauses. In *Proceedings of the 2001 ACL Workshop on Temporal and Spatial Information Processing*.
- S. Glasbey. 1993. A Computational Treatment of Sentence-Final 'then'. In *Proceedings of EACL*.
- G. Katz and F. Arosio. 2001. The Annotation of Temporal Information In Natural Language Sentences. In *Proceedings of the 2001 ACL Workshop on Temporal and Spatial Information Processing*.
- K. Krippendorff. 1980. *Content Analysis: An Introduction to its Methodology*. Sage Publications.
- I. Mani and G. Wilson. 2000. Robust temporal processing of news. In *Proceedings of the 38th Annual Meeting of the Association for Computational Linguistics (ACL 2000)*, pages 69 – 76.
- R. Mitkov. 2002. *Anaphora resolution*. Longman, London.
- MUC-6. 1995. Named entity task definition. <http://www.cs.nyu.edu/cs/faculty/grishman/muc6.html>.
- G. Puscasu. 2004. A Framework for Temporal Resolution. In *Proceedings of the 4th Conference on Language Resources and Evaluation (LREC2004)*.
- R. Quirk, S. Greenbaum, G. Leech, and J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman.
- Reuters Corpus. 2000. Volume 1, English language, 1996-08-20 to 1997-08-19, release date 2000-11-03, Format version 1. <http://about.reuters.com/researchandstandards/corpus/>.
- F. Schilder and C. Habel. 2001. From Temporal Expressions to Temporal Information: Semantic Tagging of News Messages. In *Proceedings of the 2001 ACL Workshop on Temporal and Spatial Information Processing*.
- A. Setzer and R. Gaizauskas. 2000. Annotating events and temporal information in newswire texts. In *Proceedings of LREC 2000*.
- S. Siegel and N. J. Castellan. 1988. *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, second edition edition.
- P. Tapanainen and T. Jaervinen. 1997. A non-projective dependency parser. In *Proceedings of the 5th Conference of Applied Natural Language Processing, ACL*.
- TERN. 2004. Time Expression Recognition and Normalization. <http://timex2.mitre.org/tern.html>.
- S. Weiss and C. Kulikowski. 1991. *Computer Systems That Learn*. CA: Morgan Kaufmann.