Interacting Semantic Layers of Annotation in SoNaR, a Reference Corpus of Contemporary Written Dutch

Ineke Schuurman[1] **, Véronique Hoste**[2,3]**, Paola Monachesi**[4]

[1]Centrum voor Computerlinguïstiek, K.U.Leuven
[2]LT3, Language and Translation Technology Team, University College Ghent
[3]Department of Applied Mathematics and Computer Science, Ghent University
[4]Uil-OTS, Utrecht University

Abstract

This paper reports on the annotation of a corpus of 1 million words with four semantic annotation layers, including named entities, co-reference relations, semantic roles and spatial and temporal expressions. These semantic annotation layers can benefit from the manually verified part of speech tagging, lemmatization and syntactic analysis (dependency tree) information layers which resulted from an earlier project (Van Noord et al., 2006) and will thus result in a deeply syntactically and semantically annotated corpus. This annotation effort is carried out in the framework of a larger project which aims at the collection of a 500-million word corpus of contemporary Dutch, covering the variants used in the Netherlands and Flanders, the Dutch speaking part of Belgium. All the annotation schemes used were (co-)developed by the authors within the Flemish-Dutch STEVIN-programme as no previous schemes for Dutch were available. They were created taking into account standards (either *de facto* or official (like ISO)) used elsewhere.

1. Introduction

Corpora annotated with deep semantic annotations are an indispensable resource for NLP applications such as named entity recognition, coreference resolution, semantic role labeling, etc. Due to the lack of such resources for Dutch, the transnational STEVIN-programme[1] was set up to fund a basic infrastructure which would allow researchers in linguistics and computational linguistics to perform corpus-based research. Within this STEVIN-programme, the SoNaR corpus, a 500 million words reference corpus of contemporary written Dutch is currently under construction (see also Reynaert et al. (2010)).

In this SoNaR-corpus, a core corpus of 1 million words is enriched with several layers of semantic annotation (Schuurman et al., 2009), which are manually corrected. Part of speech tagging, lemmatization and syntactic analysis for the same 1 million words has already been performed and manually corrected (Van Noord et al., 2006). While in the SoNaR-project several other tasks are performed, we concentrate in this paper on the semantic layers of the SoNaR core, and more specifically on the following four annotation layers:

- named entities (NE),

- co-reference relations (CR),

- semantic roles (SR), and

- spatial and temporal expressions (STEx).

For each of these annotation layers, the goal was to maximally reuse existing guidelines and tools, which was feasible for three out of four annotation layers. The guidelines and tools for handling co-reference in Dutch were developed by Hoste and Daelemans (2004) and further

extended in Bouma et al. (2007). The guidelines for the annotation of semantic roles and the tools for semantic role labeling are described by Monachesi, Stevens and Trapman (2007), whereas those for the annotation and detection of spatial and termporal expressions are described by Schuurman (2007a). For the named entity labeling, we did not rely on any pre-existing annotation scheme.

In the remainder of this paper, we explain each of the semantic annotation layers in a separate section. In this description, we pay special attention to the interaction between the different layers. In Section 6., we present a simplified example showing the four layers of semantic annotation, in addition to the syntactic tree annotation.[2] Finally, we describe how we plan to add even more semantic annotation layers in the future.

2. Named Entities

The annotation of named entities and the development of Dutch named entity classifiers is an under-researched area for Dutch. A notable exception is the CoNLL-2002 shared task on language-independent named entity recognition (Tjong Kim Sang, 2002), for which as small Dutch newspaper corpus was annotated with four named entity types (viz. person, organisation, location and miscellaneous). In the SoNaR project, however, we aimed at a finer granularity of the named entities and we also wanted to differentiate between the literal and metonymic use of the entities. For the development of the guidelines, we took into account the annotation schemes developed in the ACE (Doddington et al., 2004) and MUC (e.g. Chinchor and Robinson (1998)) programs, and the work on metonymy from Markert and Nissim (2002). This has led to the hierarchical annotation scheme shown in Figure 1, and of which some annotation examples are given below.

---

[1]http://taalunieversum.org/taal.technologie/stevin

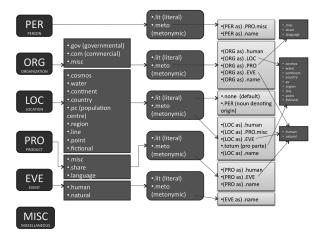[2]In reality, everything is represented in xml.

Figure 1: Named entity annotation scheme

(1) *Duitsland* (`LOC.country.lit.none`) ligt meer naar het oosten.

(English: *Germany* is located more to the East.)

(2) *Duitsland* (`LOC.country.meto.human`) won de finale.

(English: *Germany* won the finals.)

(3) Zij hebben hun zoon gisteren in *Amsterdam* (`LOC.pc.lit.none`) gezien.[3]

(English: Yesterday, they saw their son in *Amsterdam*.)

The examples clearly show that all tags consist of four parts, in which the first part of the tag denotes the `main type` of the NE, the second part the `sub type`, the third one the `use`, and the last one the `type of use`. The annotation in (1) means that *Duitsland* (Germany) is a location and more specifically a country, that it is used in literal mode. In sentence (2), on the other hand, this location is used as a metonym, representing *Duitsland* as a person. This personification of locations names, or the use of location names to refer to an associated event, is a very productive metonymic use and has been described in several studies (see for example Lakoff and Johnson (1980), Fass (1997) and Markert and Nissim (2002)). Example sentence (3) is very similar to example (1); the 'pc' refers to a population centre.

The named entity annotations are performed on raw text and were done in the MMAX2[4] annotation environment. In order to evaluate the annotation guidelines, two annotators labeled eight randomly selected texts from the corpus (about 14,000 tokens). The interannotator agreement was measured with two evaluation metrics, namely Kappa (Carletta, 1996) and F-measure ($\beta = 1$) (van Rijsbergen, 1979). The latter scores were calculated by taking one annotator as gold standard. The scores were calculated on five levels, being NE/no NE and the four NE layers. For each of the

___

[3]This sentence will be annotated for all semantic layers separately in the respective sections, while in section 6. the combination will be shown.

[4]mmax2.sourceforge.net

levels, high agreement scores were obtained, with a Kappa score ranging from 0.97 to 0.91 and an F-score ranging from 99.6% to 98.9%. For a detailled description of the guidelines and the interannotator agreement on each of the annotation levels, we refer to Desmet and Hoste (2010).

The annotated corpus is used for the development of a NE classifier (Desmet and Hoste, 2010), which will be applied to the full 500 million word SoNaR corpus. For the other three levels of semantic annotation, only the core corpus is handled.

3. Co-reference relations

The annotation of the co-reference relations started once a substantial part of NER-annotation was completed, since co-coreference annotation to a certain extent relies on named entity recognition:

(4) Minister *Keulen* werd in *Keulen* verwacht.

(English: Minister Keulen was expected in Cologne)

(5) *Den Haag* speelt dit weekend thuis, in hun nagelnieuwe stadion in *Den Haag*

(English: This weekend Den Haag plays a home match, in their brand new stadium in The Hague)

In 4 and 5 both *Keulen* and *Den Haag* are `not` used in a co-referential relation, which can be deduced automatically from the NE annotations. In example (4), the first occurrence of *Keulen* is annotated as (`PER.lit.name`), whereas the second occurrence receives the label (`LOC.pc.lit.none`). The same holds for sentence (5) in which we have (`LOC.pc.meto.human`) vs (`LOC.pc.lit.none`).

For the annotation of the co-referential relations, we followed an existing annotation scheme ((Hoste and Daelemans, 2004), (Bouma et al., 2007)). The scheme differentiates between the following nine co-reference relations, of which the first relation type is by far the most frequent one:

1. Identity (or strict co-reference)
2. Time-indexed co-reference
3. Type-token co-reference
4. Part/whole co-reference
5. Modality and negation
6. Predicate nominals
7. Appositions
8. Bound anaphora
9. Metonymy

Co-reference links were annotated between nominal constituents, which could take the form of a pronominal, named entity or common noun phrase, as exemplified in (6 and 7).

(6) *Zij*[id="1"] hebben *hun*[ref="1" type="ident"] zoon gisteren in Amsterdam gezien.

(7) In de Raadsvergadering is het vertrouwen opgezegd in *het college* [id="2"]. In een motie is gevraagd aan *alle wethouders* [ref="2" type="bridge"[5]] om ontslag in te dienen.

(English: In the council meeting the confidence in mayor-and-aldermen has been withdrawn. A motion requests that all aldermen resign.)

In order to avoid conflicts between the annotation layers, the co-reference annotations were performed on the nominal constituents which were extracted from the manually validated syntactic output (Van Noord et al. 2006). The brackets in Figure 2 indicate the NPs extracted from the dependency tree output; also embedded NPs are indicated. Figure 2 shows the coreference relation annotation between the NPs "Koninkrijk, samenleving en economie" and "de drie terreinen".

(English: In my introduction, I will briefly describe the three domains - Kingdom, society and economy (...))

Figure 2: Example coreference annotation in MMAX2

Since inter-annotator agreement for this labeling task was already measured in the framework of the design of the annotation guidelines (Hendrickx et al., 2008), no separate inter-annotator agreement assessment was done. Hendrickx et al. (2008) computed the inter-annotator agreement on the identity relations as the F-measure of the MUC-scores (Vilain et al., 1995) obtained by taking one annotation as 'gold standard' and the other as 'system output'. They report an inter-annotator agreement of 76% F-score on the identity relations. For the bridging relations, an agreement of 33% was reported.

Due to the low performance of the current classification-based coreference resolution systems for Dutch (Hoste, 2005; Hendrickx et al., 2008) no automatic pre-annotation was performed to support or accelerate the annotation process. The current classification-based co-reference resolution systems for Dutch report an optimal MUC F-score of about 50% on the identification of identity relations. Once a substantial part of the corpus is annotated with

---

[5]The "bridge" label is used to annotate part/whole co-reference

co-reference information, the current classifier will be re-trained and we will then assess the usefulness of integrating this pre-annotation.

4. Semantic role labeling

Semantic role labeling takes into account the results of the previous layers, the annotation of a small fragment consisting of about 3000 predicates has already been completed.

For the development of the guidelines, we have considered the annotation scheme proposed within existing projects such as FrameNet (Johnson et al., 2002) and PropBank (Kingsbury et al., 2002). However, we have decided to adapt the PropBank annotation scheme because of the promising results with respect to automatic semantic role labeling which have been obtained for English.

The guidelines for Dutch were developed by Monachesi, Stevens and Trapman (2007). More specifically, the Prop-Bank annotation scheme needed to be revised due to the different theoretical approach adopted within Sonar. In the case of traces, PropBank creates co-reference chains for empty categories while in our case, empty categories are almost non existent and in those few cases in which they are attested, a coindexation has been established already at the syntactic level. Furthermore, in Sonar we assume dependency structures for the syntactic representation while PropBank employs phrase structure trees. In addition, Dutch behaves differently from English with respect to certain constructions (i.e. middle verb constructions) and these differences needed to be spelled out.

Besides the adaptation (and extension) of the guidelines to Dutch, we also needed a Dutch version of the PropBank frameindex. In PropBank, frame files provide a verb specific description of all possible semantic roles and illustrate these roles by examples. The lack of example sentences makes consistent annotation difficult. Since defining a set of frame files from scratch is very time consuming, we have annotated Dutch verbs with the same argument structure as their English counterparts, thus using English frame files instead of creating Dutch ones. This approach proved to be successful in the majority of the cases.

During the adaptation, several problems emerged which need to be solved in the SoNaR project. These concern both linguistic issues as well as issues related to the interaction with previous levels of annotation. For example, the interpretation of modifiers is a complex phenomenon for which linguistic research doesn't provide a standard solution yet and for which we should provide a possible annotation.

Even though the syntactic corpus which has been employed for the annotation of the semantic roles had been manually corrected, we have encountered examples in which the annotation provided by the syntactic parser was not appropriate. This is the case of a PP which was labeled as modifier by the syntactic parser but which should be labeled as argument according to the PropBank guidelines.

Furthermore, we have encountered problems with respect to PP attachment, that is the syntactic representation gives us correct and incorrect structures and at the semantic level we are able to disambiguate.

It should be noted that at this stage, annotation has been carried out by only one annotator, therefore we have not

been in the condition to measure inter-annotator agreement. An example of sentence annotated with PropBank roles can be found in the example below:

(8) *Zij* [ARG0] hebben *hun zoon* [[ARG2] *gisteren* [[ARGM-TMP] *in Amsterdam* [[ARGM-LOC] *gezien* [[PRED]

One advantage of employing PropBank for the annotation of semantic roles is that it is quite suitable for `automatic` semantic role labeling. However, in the case of Dutch, there was no semantically annotated corpus available that could be used as training data. For the D-Coi corpus, a novel approach to rule-based tagging based on D-Coi dependency trees has been proposed (Stevens, 2006). A semantic argument tagger, called XARA (XML-based Automatic Role-labeler for Alpino-trees) has been developed. It establishes a basic mapping between nodes in a dependency graph and PropBank roles. A rule in XARA consist of an XPath expression that addresses a node in the dependency tree, and a target label for that node, i.e. a rule is a *(path,label)* pair. The evaluation carried out shows that XARA achieves a precision of 65,11%, a recall of 45,83% and an F-score of 53.80.

The output of the corpus automatically annotated by means of XARA has been manually corrected and it has been used as training and test data for an SRL classification system. For this learning system we have employed a Memory Based Learning (MBL) approach, implemented in the Tilburg Memory based learner (TiMBL).[6] The classifier obtained a precision of 70.27%, a recall of 70.59% and an F-score of 70.43. For SoNaR, the classifier is further adapted, taking into account the results of the new annotation layers of NER, and especially Co-reference.

5. Spatial and temporal relations

Whereas usually these two layers of annotation are handled separately, we will be using STEx, a combined spatiotemporal annotation scheme. This scheme takes into account both TimeML (TimeML Working Group (2010) as well as Sauri et al. (2006), upon which the ISO standard ISO TimeML is mainly based) and SpatialML (2007), an ISO standard under construction. A first version of STEx, MiniSTEx, was developed within the D-Coi project, the tool used there being a semi-automatic one.
Within the SoNaR project, the STEx spatial scheme will be largely restricted to geospatial annotation.[7]
Like MiniSTEx (Schuurman, 2007a; Schuurman, 2007b; Schuurman, 2008), the current system STEx handles spatial and temporal expressions basically the same way, cf. Table 1 below.
Besides the fact that STEx uses geospatial information to determine temporal information and the other way around, STEx also differs from both TimeML and SpatialML in that it is provides more details, cf. Schuurman (2007b), Schuurman (2008).

temporal	geospatial
time of perspective	place of perspective
time of location	place of location
time of eventuality	place of eventuality
duration	distance
shift of perspective	shift of perspective
relations	relations

Table 1: The resemblance between temporal and spatial analyses

(9) Zij hebben hun zoon *gisteren* [temp type="cal" ti="tp-1" unit="day" val="2008-05-22"] in *Amsterdam* [geo type="place" val="EU::NL::-::NH::Amsterdam::Amsterdam" coord="52.3666666666667,4.9"] *gezien* [temp type="event" value="vtt" rel="before(ti,tp)"]

In (9) the time-zone associated with it (time-zone="UTF+1") is filtered out, although it is contained in the metadata coming with the text. Only when its value is overruled by a statement in the text it will be mentioned in the annotation itself. In the full xml-format it is represented everywhere, i.e. also when it is not overruled. (9) also contains a shorthand version of the formulas we associated with several temporal expressions. `ti="tp-1"` `unit="day"` says that the *time of eventuality* ti is the *time of perspective* tp minus 1. As the `unit` involved is that of *day*, only that variable is to be taken into account. So, *yesterday* is to be associated with a formula, not with an accidental value (like "2008-05-22" in (9)). In a second step, the calculations are to be performed. This is not only of importance for our current, mainly rule-based system, but, at a later stage, also for a machine learning approach.
In the framework of the SoNaR corpus, STEx makes use of the information available through previous syntactic and semantic layers.[8] In some cases it completes and disambiguates such information. For example, the location related annotations at the level of NER will be disambiguated. When a sentence like (9) occurs in a document, usually an expression like *Amsterdam* can be disambiguated, stating that the instantiation of *Amsterdam* meant is the town of Amsterdam in the Netherlands, not one of the towns or villages in the US, Canada, ... Especially in a corpus, the metadata coming with a file allow for such an annotation, cf (Schuurman, 2007b). Co-reference will also be very useful, the same holds especially for metonymy as annotated in NER (see also (Leveling and Hartrumpf, 2008))
Note that for example with respect to SRL, there sometimes is a mismatch, due to different definitions. Something labeled ARG-LOC is not necessarily a spatial item at the level of spatiotemporal annotation, as in SRL abstract locations are also taken into account (*'In his speech he told us ...'*). Still, in many cases, information with respect to SR

---

[6]http://ilk.uvt.nl/timbl/
[7]As is the case in SpatialML as well.

[8]In another project, AMASS++ (cf. (Schuurman and Vandeghinste, 2010)) a version of STEx is being used in which it has to rely on automatic PoS tagging and chunking. In a later paper we intend to compare such approaches: is manual correction/addition of further layers of annotation worth the effort (time and money)?

is helpful, cf. Llorens et al. (2009).

Spatiotemporal annotation in SoNaR is performed automatically, using a large data base containing geospatial and temporal data, combinations of these and especially also cultural data with respect to such geospatial and temporal data. For example: what is considered as the begin and end dates of World War II is not the same all over Europe and the rest of the world. The same holds for the date(s) associated with Christmas, or Thanksgiving. Or to decide which *Cambridge* (UK, US) is refered to, or which *Antwerpen* (Antwerp): the province, the municipality or the populated place.[9] *Cultural* aspects like tradition (Jewish, Christian), geographical background, social background have their effects on the (intended) interpretation of temporal and geospatial data, cf. Figure 3.
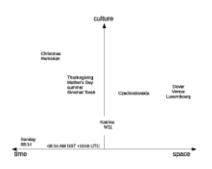

Figure 3: Eventualities with temporal, geopatial or and/or cultural aspects

Each annotation is in principle corrected by one corrector (student), substantial parts are corrected by more students in order to ensure annotator agreement. Work for the spatiotemporal layer of annotation within the SoNaR project is expected to start summer 2010, as substantial parts of the corpus annotated (and manually corrected) with the other semantic layers should be available as input.

6. Resulting annotation

For human users, in-line annotations as shown in the previous sections give a clear picture for the named entities, co-reference relations and semantic roles, even when they are represented in xml-format. The spatio-temporal relations, however, are less readable. That also becomes the case when a user is interested in a combination of, for example, NE and CR, or SR in combination with the original tree showing syntactic dependencies. A tree-like representation, such as shown in Figure 4, provides more insight in these cases. Such trees can be evoked by users for all layers, alone or in combination, on basis of the (combined) standoff xml-annotations. Note that these trees are sentence-based and that links with previous sentences are not shown.

This problem can be overcome by the use of so-called megatrees, cf. Mladová et al. (2008). In such megatrees several sentence-level trees are combined, thus covering whole paragraphs. In order to promote their readability,

we will implement a feature allowing to hide the content of a node temporarily by clicking on it. After a second click the content will pop up again. Thus, users (amongst them the students who are to correct the annotations: xml-files are not suitable for manual correction) can click away all those (parts of) trees they are not interested in at a particular moment.

Such megatrees are also great for comparison of the various layers of annotation. When do you have an ARG-LOC that is not considered a spatial expression at the level of STEx or a locative at the level of NE? Such 'contradictions' are perfectly possible because the various layers of annotation were developed independently, each building on official or *de facto* standards on hand ar the time of development. Structural deviations therefore are to be expected. Accidental deviations, on the other hand, are prone to indicate mistakes.

7. More plans for the future

In a recently (March 2010) started project *TST Tools voor het Nederlands als Webservices in een Workflow* (TTNWW, HLT Tools for Dutch as Web Services in a Work Flow), a Flemish-Dutch CLARIN pilot financed by the Flemish (via EWI) and Dutch (via CLARIN-NL) governments, the semantic annotations discussed in this paper will easily be made accessible, especially for researchers in human and social sciences.

We also plan to add more layers of annotation, like sentiment and emotion analysis and discourse analysis.

8. Acknowledgements

9. References

G. Bouma, W. Daelemans, I. Hendrickx, V. Hoste, and A. Mineur. 2007. The COREA-project, Manual for the annotation of coreference in Dutch texts. Technical report, University Groningen.

J. C. Carletta. 1996. Assessing agreement on classification tasks: the kappa statistic. *Computational Linguistics*, 22(2):249–254.

N. Chinchor and P. Robinson. 1998. MUC-7 Named Entity Task Definition (version 3.5).

B. Desmet and V. Hoste. 2010. Toward a Balanced Named Entity Corpus for Dutch. Proceedings of the seventh international conference on Language Resources and Evaluation.

G. Doddington, A. Mitchell, M. Przybocki, R. Ramshaw, S. Strassel, and R. Weischedel. 2004. The Automatic Content Extraction (ACE) Program Tasks, Data, and Evaluation. In *Proceedings of LREC 2004*.

D. Fass. 1997. *Processing Metaphor and Metonymy*. Ablex, Stanford, CA.

I. Hendrickx, V Hoste, and W. Daelemans. 2008. Semantic and Syntactic features for Anaphora Resolution

---

[9]At the moment, the precision for such geospatial anchors in STEx is 0.92, recall 0.91 (small scale test for some 200 instances).

top

smain
temp   event
       tp=2008-05-23
       vtt   before(ti,tp)
       ti=2008-05-22
geo   in(i,Amsterdam)

punct
.

su:1   hd   vc
VNW(pers,pron,nomin,vol,3p,mv)   WW(pv,tgw,mv)   ppart
role   ARG0   hebben
coref   id="1"
zij

su:1   obj1   mod   ld   hd
            NP   BW()   PP   role   PRED
            role   ARG2   role   ARGM-TMP   role   ARGM-LOC gezien
                            temp   cal
                                   ti=tp-1
                                   day
                                   2008-05-22
                            NE   date
                                   gisteren

det   hd   hd   obj1
VNW(bez,...,3,mv,prenom,...) N(soort,ev,basis,zijd,stan) VZ(init)   N(eigen,ev,basis,onz,stan)
coref ref="1"   zoon   in   geo   place
hun   NL::-::NH::A'dam::A'dam
                                                        NE   location
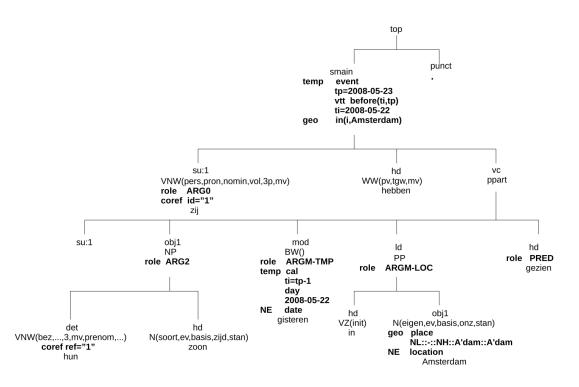   Amsterdam

Figure 4: An enriched tree (simplified)

for Dutch. In *Lecture Notes in Computer Science, Volume 4919, Proceedings of the CICLing-2008 conference*, pages 351–361. Berlin: Springer Verlag.

V. Hoste and W. Daelemans. 2004. Learning Dutch Coreference Resolution. In *Computational Linguistics in The Netherlands 2004*, pages 133–148.

V. Hoste. 2005. *Optimization Issues in Machine Learning of Coreference Resolution.* Ph.D. thesis, Antwerp University.

G. Lakoff and M. Johnson. 1980. *Metaphors we live by.* Chicago: University of Chicago.

J. Leveling and S. Hartrumpf. 2008. On Metonymy Recognition for Geographic Information Retrieval. *International Journal of Geographical Information Science*, 22(3):289–299.

H. Llorens, B. Navarro, and E. Saquete. 2009. Using Semantic Networks to Identify Temporal Expressions from Semantic Roles. In *Proceedings of RANLP-2009*.

K. Markert and M. Nissim. 2002. Towards a Corpus Annotated for Metonymies: the Case of Location Names. In *Proceedings of the Third International Conference on Language Resources and Evaluation*, pages 1385–1392.

L Mladová, Š. Zikánová, and E. Hajičová. 2008. From Sentences to Discourse: Building an Annotation Scheme for Discourse Based on Prague Dependency Treebank. In *Proceedings of LREC*.

P. Monachesi, G. Stevens, and J. Trapman. 2007. Adding semantic role annotation to a corpus of written Dutch. In *Proceedings of LAW-07*, Prague. Czech Republic. ACL 2007 workshop.

M. Reynaert, N. Oostdijk, O. De Clercq, H. van den Heuvel, and F. de Jong. 2010. Balancing SoNaR: IPR versus processing issues in a 500-million-word written Dutch reference corpus. In *Proceedings of LREC*, Malta.

R. Sauri, J. Littman, B. Knippen, R. Gaizauskas, A. Setzer, and J. Pustejovsky, 2006. *TimeML Annotation Guidelines, version 1.2.1.*

I. Schuurman and V. Vandeghinste. 2010. Cultural Aspects of Spatiotemporal Analysis in Multilingual Applications. In *Proceedings of LREC*.

I. Schuurman, V. Hoste, and P. Monachesi. 2009. Cultivating Trees: Adding Several Semantic Layers to the Lassy Treebank in SoNaR. In *Proceedings of the 7th International Workshop on Treebanks and Linguistic Theories*.

I. Schuurman. 2007a. Spatiotemporal Annotation on Top of an Existing Treebank. In K. De Smedt, J. Hajic, and S. Kuebler, editors, *Proceedings of the Sixth International Workshop on Treebanks and Â Linguistic Theories*, pages 151–162, Bergen, Norway.

I. Schuurman. 2007b. Which New York, which Monday? The role of background knowledge and intended audience in automatic disambiguation of spatiotemporal expressions. In *Proceedings of CLIN 17*.

I. Schuurman. 2008. Spatiotemporal annotation using MiniSTEx: How to deal with alternative, foreign and obsolete names? In *Proceedings of LREC 2008*, Marrakech, Morocco.

SpatialML. 2007. Annotation Scheme for Marking Spatial Expressions in Natural Language, October 1. MITRE Corporation.

G. Stevens. 2006. Automatic Role Labeling in a Dutch Corpus. Master's thesis, Utrecht University.

TimeML Working Group, 2010. *TimeML Annotation Guidelines, version 1.3*, February 9.

E.F. Tjong Kim Sang. 2002. Memory-based Named Entity Recognition. In *Proceedings of CoNLL-2002*, pages 203–206.

G. Van Noord, I. Schuurman, and V. Vandeghinste. 2006.

Syntactic Annotation of Large Corpora in STEVIN. In *Proceedings of LREC 2006*, Genua.

C.J. van Rijsbergen. 1979. *Information Retrieval*. Buttersworth, London.

M. Vilain, J. Burger, J. Aberdeen, D. Connolly, and L. Hirschman. 1995. A Model-Theoretic Coreference Scoring Scheme. In *Proceedings of the Sixth Message Understanding Conference (MUC-6)*, pages 45–52.