

Open Research Online

The Open University's repository of research publications and other research outputs

A short survey of discourse representation models

Conference or Workshop Item

How to cite:

Groza, T.; Handschuh, S.; Clark, T.; Buckingham Shum, S. and de Waard, A. (2009). A short survey of discourse representation models. In: Proceedings 8th International Semantic Web Conference, Workshop on Semantic Web Applications in Scientific Discourse. Lecture Notes in Computer Science, Springer Verlag: Berlin, 26 Oct 2009, Washington DC.

For guidance on citations see [FAQs](#).

© 2009 The Authors

Version: Accepted Manuscript

Link(s) to article on publisher's website:
<http://ceur-ws.org/Vol-523>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's [data policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

A Short Survey of Discourse Representation Models

Tudor Groza¹, Siegfried Handschuh¹, Tim Clark², Simon Buckingham Shum³, and
Anita de Waard⁴

¹ DERI, National University of Ireland, Galway,
IDA Business Park, Lower Dangan, Galway, Ireland
{tudor.groza, siegfried.handschuh}@deri.org

² Initiative in Innovative Computing, Harvard University
60 Oxford Street, Cambridge, MA 02138, USA
twclark@nmr.mgh.harvard.edu

³ Knowledge Media Institute, The Open University
Milton Keynes, MK7 6AA, UK
S.Buckingham.Shum@open.ac.uk

⁴ Elsevier Labs, Radarweg 29, 1043 NX
Amsterdam, The Netherlands
A.dewaard@elsevier.com

Abstract. With the advancement of technology and the wide adoption of ontologies as knowledge representation formats, in the last decade, a handful of models were proposed for the externalization of the rhetoric and argumentation captured within scientific publications. Conceptually, most of these models share a similar representation form of the scientific publication, i.e. as a series of interconnected elementary knowledge items. The main differences are given by the terminology used, the types of rhetorical and / or argumentation relations connecting the knowledge items and the foundational theories supporting these relations. This paper analyzes the state of the art and provides a concise comparative overview of the five most prominent discourse representation models, with the goal of sketching an unified model for discourse representation.

1 Introduction

Dissemination, an important phase of scientific research, can be seen as a communication process between scientists. They expose and support their findings, while discussing claims stated in related scientific publications. This communication takes place over the course of several publications, where each paper itself contains a rhetorical discourse structure laying out supportive evidence for the raised claims. This discourse structure is usually hidden in the semantics expressed by the writer within the publication's content and thus hard to discover by the reader.

Externalization, as defined by Nonaka [1], represents *the process of articulating tacit knowledge into explicit concepts*. It holds the key to social knowledge creation and crystallization, through a process of sharing, discussion and testing with others. In the scientific dissemination context, externalization has a dual form. On the one hand, scientific publications represent intrinsically a form of cognitive externalization, making explicit the scientists' thoughts. On the other hand, in order to make

these publications much more accessible to computation, and more specifically on the Web, so that information can be easier navigated, compared and understood, we need for a formal externalization, i.e. stepping from the freely expressed text to machine-processable structures. In the case of the rhetorical and argumentation discourse based on claims and evidence, the degree of formalization can be a couple of keywords, or a weakly structured text, both possibly including direct references to the publications first stating the actual claims or providing evidence for new claims.

In the last decade, a handful of models targeting the externalization of the rhetoric and argumentation captured within the discourse of scientific publications, were proposed. Conceptually, most of these models share a similar representation form of the scientific publication, i.e. as a series of interconnected elementary knowledge items. The main differences are given by the terminology used, the types of rhetorical and / or argumentation relations connecting the knowledge items and the foundational theories supporting these relations. While the argumentation side is, in general, inspired from IBIS (Issue Based Information Systems) [2], the rhetorical structuring uses different foundations, such as Cognitive Coherence Relations [3] or the Rhetorical Structure Theory (RST) [4]. A number of discourse-relationship approaches, including a comparison between RST and other taxonomies has been discussed by Hovy [5].

This paper provides a concise comparative overview of some of these discourse representation models. We will focus on five such models: SWAN (Semantic Web Applications in Neuromedicine) [6], SALT (Semantically Annotated L^AT_EX) [7] and the models proposed by the Scholarly Ontologies project [8], Harmsze [9] and de Waard [10]. We would like to point out that in addition to these representation-driven approaches, and in the same category of analysis of scientific publications content, research was also performed on automatic extraction of epistemic items. Relevant work includes the efforts of Teufel [11], Mizuka [12] or Lisacek [13]. Nevertheless, in this paper, we concentrate only on the former approaches with the goal of trying to find a common denominator that could lead to an unified model for discourse representation.

The remainder of the paper is structured as follows: Sect. 2 lists the aspects on which we have focused for the comparative analysis, Sect. 3 details the five above-mentioned models and before concluding in Sect. 5, we discuss the overall comparison of the representations in Sect. 4.

2 Analysis features

For building a comprehensive comparison of the discourse representation models, we compiled a list of features to be observed throughout the overall analysis. The list consists of the following elements:

- **Course-grained rhetorical structure** – identifies the existence of a course-grained rhetorical structure representation within the model. Its goal is to capture the semantics of larger blocks of text inside the publication’s content that have an associated rhetorical role.
- **Fine-grained rhetorical structure** – as opposed to the previous feature, this feature considers the fine-grained content composing the discourse (i.e. restricted discourse knowledge items in forms of claims, positions, arguments, etc)

between which usually emerges a network arrangement driven by the different types of relations that connect the fine-grained elementary items.

- **Relations** – looks at the types of relations used for linking the fine-grained structure into an unitary network.
- **Polarity** – specifies if the model includes explicitly the polarity of the relations (i.e. positive or negative). For example, a *supports* relation would have a positive polarity attached, while a *refutes* relation would have a negative polarity. Generally, this polarity is to some extent similar to the polarity extracted in the opinion mining and sentiment analysis field, which, we will not focus on, since it is out of the scope of this paper.
- **Weights** – specifies if the model considers explicitly the weights of the relations, i.e. if some relations are stronger than others. This feature can be tightly coupled to the polarity. For example, the *supports* relation might be considered stronger than the *agrees with* relation, both being positive from the polarity perspective.
- **Provenance** – indicates whether the model encapsulates also the provenance information attached to the fine-grained rhetorical structure (i.e. the accurate localization of the text span that represents the textual counterpart of the discourse knowledge item).
- **Shallow metadata support** – shows if the model has embedded support for shallow metadata (e.g. authors, titles, etc)
- **Domain knowledge** – analyses the close coupling of the model to particular domain knowledge areas.
- **Purpose** – presents the purpose, or intended use, of the model as envisioned by their creators.
- **Evaluation and uptake** – mentions the evaluation and uptake status of the model.

These last two features in the list try to capture the “practicality” dimension of the discourse representation models, with the last one pointing in essence to a reality-check, in terms of deployment, adoption and adequacy of the models in actual use by scientists.

3 Discourse representation models

3.1 Harmsze’s Model

One of the first and probably the most comprehensive models for capturing the rhetoric and argumentation within scientific publications was introduced by Harmsze [9]. She focused on developing a modular representation for the creation and evaluation of scientific articles. Although the corpus used as a foundation for the analysis was about experimental molecular dynamics, the resulted model is uniformly valid for any scientific domain.

The author models the discourse by means of a coarse-grained structure split into modules and a series of links to connect these modules. The six modules proposed by Harmsze are as follows: (i) **Meta-Information** is a support module that keeps the entire publication glued together. It consists of several parts, such as, the bibliographic information, abstract, lists of references or acknowledgements; (ii) **Positioning** sets the context of the research presented in the publication. It describes the *situation* in which the research issues are considered and the *central problem* of

the research. (iii) **Methods** acts as a container for the authors' response to the central problem. The model provides three types of possible methods, i.e. experimental, numerical and theoretical methods. (iv) **Results** details the results achieved with the methods previously mentioned. It consists of *raw data* and the *treated results*. (v) **Interpretation** contains the authors' interpretation of the results. It usually deals with the process of interpreting the results and the argumentation of the plausibility and on the relevance of the interpretation. (vi) **Outcome** aggregates the authors' *findings* and the *leads to further research*.

To connect the above mentioned modules, the model introduces two types of relations: (i) organizational links, and (ii) scientific discourse relations. The *organizational links* provide the reader with the means to easily navigate between the modules composing the scientific publication. They connect only modules as entire entities and do not refer to the segments encapsulated in them, which in turn would identify the content. Harmsze distinguishes six types of organizational links: *hierarchical*, *proximity*, *range-based*, *administrative*, *sequential* and *representational*. On the other hand, regarding the links between segments of modules (scientific discourse relations), the model describes two main categories: *relations based on the communicative function*, that have the goal of increasing the reader's understanding and maybe acceptance of the publication's content, and *content relations*, that allow the structuring of the information flow within the publication's content. The first category is split into: *Elucidation*, as *Clarification* and *Explanation*, and *Argumentation*. The second category contains: *Dependency in the problem-solving process*, *Elaboration*, as *Resolution* and *Context*, *Similarity*, *Synthesis*, as *Generalization* and *Aggregation*, and *Causality*. Generally, the relations present an implicit polarity and don't have attached explicit weights or temporal aspects.

From the evaluation perspective, the authors performed a preliminary evaluation of the model, which showed that the model satisfies the purpose for what it was designed, but in reality, to our knowledge, it was not deployed in an actual application and consequently it failed to be adopted.

3.2 The Scholarly Ontologies project

A much more focused approach was the one followed by Buckingham Shum et al. [8] in the Scholarly Ontologies (ScholOnto) project. They were the first to propose the decomposition of a scientific publication into elementary discourse knowledge items and their connection via a set of relations emerged from an established theoretical foundation, i.e. Cognitive Coherence Relations [3]. As opposed to other representations, Buckingham Shum et al. do not model the coarse-grained rhetorical or linear structure of the publications, but rather concentrate strictly on organizing the coherence among the content segments. Their research resulted in a series of tools for the annotation and visualization of argumentation (the latest with accent on Web 2.0 technologies [14]), that acted as inspiration to other approaches.

The elementary discourse knowledge items introduced by Buckingham Shum's model are the *atomic nodes*, that represent short pieces of text, within the publication, succinctly summarizing the authors' contribution. The granularity of these nodes is left for decision to the author, and thus, can vary from parts of sentences to blocks of sentences. Nodes can have several types (e.g. Data, Language, Theory), encoded in the links that connects them. Two such connected nodes form a *Claim*. In addition to nodes, the model contains also two kinds of composite elements:

(i) **sets** that group several nodes sharing a common type (or theme), and (ii) **claim triples** formed by linking sets or atomic nodes.

In terms of relations, Buckingham Shum’s *Discourse Ontology* comprises six main types: (i) *causal* links, e.g. predicts, envisages, causes or prevents; (ii) *problem related* links, e.g. addresses or solves; (iii) *similarity* links, e.g. is identical to, is similar to, or shares issues with; (iv) *general* links, e.g. is about, improves on, or impairs; (v) *supports / challenges* links, e.g. proves, refutes, is evidence for, or agrees with; (vi) *taxonomic* links, e.g. part of, example of, or subclass of Each relation has attached an explicit polarity (positive or negative), and a specific weight. The link’s polarity denotes explicitly the author’s position in regards to particular statements present in the related work, similar to Teufel’s approach [15]. At the same time, the weight indicates how strong or weak is the author’s position. For example, the *causal* links *envisages* and *causes* have both a positive polarity, but different weights, the former being considered weaker than the latter. Similarly, *is unlikely to affect* and *prevents* have a negative polarity with different weights, again the latter being considered stronger than the former.

The authors performed an extensive evaluation of their approach, the model being deployed in several applications developed by the authors, such as, Compendium⁵ and Cohere⁶, applications that are currently widely used.

3.3 De Waard’s Model

A different discourse representation model was proposed by de Waard [10]. They started with a rhetorical block structure for scientific publications called ABCDE, similar to the IMRAD (Introduction, Material and Methods, Results and Discussion)⁷ structure. The title holds the acronym of the five types of blocks present in the model: (i) *Annotations*, representing the set of shallow metadata associated with each publication (usually expressed in DublinCore⁸ terms) (ii) *Background*, describing the positioning of the current research and the ongoing issues; (iii) *Contribution*, describing the work performed by the authors; (iv) *Discussion*, comparing the current work to other approaches, including implications and next steps; (v) *Entities*, denoting references, personal names, project websites, etc.

At a later stage, the authors enriched their model with a fine-grained representation of the discourse, by identifying seven basic discourse segment types [A]: Fact, Hypothesis, Goal, Method, Result, Implication, and Problem. These types correspond to those found by Mizuka et al.[12] using automated techniques based on the argumentation zoning approach developed by Teufel et al.[15] in a corpus of biology texts. A first attempt was made to find these segment types computationally [16].

3.4 The SWAN Ontology

The SWAN (Semantic Web Applications in Neuromedicine)⁹ project focuses on developing a semantically structured framework for representing biomedical discourse.

⁵ <http://compendium.open.ac.uk/institute/>

⁶ <http://cohere.open.ac.uk/>

⁷ <http://www.uio.no/studier/emner/hf/imk/MEVIT4725/h04/resources/imrad.xml>

⁸ <http://dublincore.org/>

⁹ <http://swan.mindinformatics.org/ontology.html>

Initially, the framework was applied to significant problems in Alzheimer Disease research, though it is not restricted only to this particular domain. The purpose of the SWAN Ontology is to function as a schema of distributed knowledge based in Alzheimer Disease and to link information intrinsically captured with information residing in other biomedicine sources.

The ontology has six main elements: (i) **people, groups and organizations**, modeled via well-known entities of the FOAF¹⁰ vocabulary, in addition to a series of properties introduced by SWAN; (ii) **bibliographic records and citations**, captured based on properties similar to the ones proposed by the Bibliographic Ontology (BIBO)¹¹, i.e., *DOI*, *ISSN*, *JournalName* or *Issue*; (iii) **life science entities**, representing domain knowledge deeply integrated within the framework, and referring to concepts such as *Gene*, *Protein*, *Organism* or *Multiprotein Complex*, together with their associated relations (e.g. *encodes*, *containsProtein* or *nativeIn*); (iv) **tags, qualifiers and vocabularies** supporting personal data organization; (v) **versions and provenance** modeling temporal and provenance information, such as creation date of an entity, update status or authorship; (vi) **discourse elements** acting as the core of SWAN and modeling the rhetorical structure of the discourse.

SWAN proposed four types of discourse elements connected via a mixture of argumentation and cognitive coherent relations with implicit polarity (e.g. *consistentWith*, *inconsistentWith*, *discusses*, *alternativeTo*, *citesAsEvidence* or *inResponseTo*): (i) **Discourse Element** is a narrative object, representing a mapping of digital resources to statements in natural language (e.g. sentences, paragraphs). Each such element may be linked dynamically to terms or statements in other domain ontologies; (ii) **Research Statement** represents a particular discourse element having a claim or hypothesis nature; (iii) **Research Question** is a topic under investigation; (iv) **Structured Comment** acts as a structure representation of a comment published in a digital resource.

Similar to the Scholarly Ontologies project, this model was also extensively evaluated and is currently in use by the Alzheimer Disease research community and under evaluation by others.

3.5 The SALT Framework

SALT (Semantically Annotated L^AT_EX)¹² [7] represents a semantic authoring framework targeting the enrichment of scientific publications with semantic metadata. SALT adopts elements from the Rhetorical Structure Theory (RST) [4] with the goal of modeling discourse knowledge items and their intrinsic coherence relations. The framework comprises three ontologies: (i) the Document Ontology, modeling the linear structure of a document, in terms of *Sections*, *Paragraphs* or *TextChunks* (ii) the Rhetorical Ontology, capturing the rhetorical and argumentation structure of the publication, and (iii) the Annotation Ontology, that connects the rhetorical structure present within the document's content to the actual content of the document. This ontology acts as a semantic bridge between the other two ontologies and in addition it re-uses well-known concepts and properties for exposing shallow metadata from the FOAF vocabulary.

¹⁰ <http://www.foaf-project.org/>

¹¹ <http://bibliontology.com/>

¹² <http://salt.semanticauthoring.org/>

The Rhetorical Ontology consists of three major sides: (i) **rhetorical relations** side that models elementary rhetorical elements (e.g. *claims* or *supports*) and the relations connecting them (e.g. *antithesis*, *circumstance*, *concession* or *purpose*); (ii) **rhetorical blocks** side that provides a coarse-grained structure for modeling the discourse (e.g. *abstract*, *motivation*, *background* or *conclusion*); (iii) **argumentation** side that captures the argumentation present in the publication via concepts like *Issue*, *Position* or *Argument*.

Applying SALT on a scientific publication leads to a local instance model capturing the inter-connected linear, rhetorical and argumentation structures within that publication. At a later stage, the authors dealt with the global scope of modeling discourse knowledge items, i.e. items and relations that span across multiple publications. In [17], following a semiotic approach inspired from Peirce’s direction of semiotics (the semiotic triangle [18]), the authors introduce a model for externalizing argumentative discourse networks.

Different aspects of SALT were evaluated in a series of experiments in the last three years, while the model as a whole was recommended for creating semantic metadata for scientific publications by different workshops, such as the SemWiki (Semantic Wiki) workshop series, or the Scripting and Development for the Semantic Web workshop series.

4 Discussion: Towards an unified discourse representation

Fig. 1 presents a concise comparative overview of the five discourse representation models we have previously described. To make the first steps towards an unified discourse representation model, we believe that we have to find a proper balance between the features each of the currently existing models presents. In the following we will try to sculpt the skeleton of such an unified model, to be later discussed within the community.

The first aspect to be considered is the overall structure of the model. By following a layered approach, such as the one proposed by SWAN and SALT, the unified model will gain flexibility, which in turn will be reflected in a more straightforward evolution. This would clearly decouple the rhetoric and argumentation from the provenance information, and from the shallow metadata and domain knowledge, while at the same time providing the opportunity for a modular enrichment of the model as a whole.

The second aspect is the discourse structuring level. To be able to capture the complete semantics hidden within the discourse, the model needs to address it at different levels. Consequently, it needs to present both a coarse-grained structure, as well as, a fine-grained structure. The former can be easily created by adopting a mixture of rhetorical blocks from de Waard’s model and SALT, or from a different source, such as Teufel’s zones [11]. These blocks would have the role of organizing the publication’s rhetoric at a high level. The latter structure refers to decomposing the publication’s content into fine-grained discourse knowledge items, to be connected via different types rhetorical and argumentation relations. This will lead to a network of inter-linked elementary items that will externalize the content’s coherence and argumentation thread. Fortunately, all the presented models contain such a fine-grained structure, the only difference being the terminology used. We believe that a core term such as *Discourse element*, with an underlying synonymy to claim or hypothesis should be easily acceptable.

Feature/ Approach	Coarse- grained rhetorical structure	Fine-grained rhetorical structure	Relations	Polarity	Weights	Provenance	Shallow Metadata Support	Domain Knowledge	Purpose	Evaluation and uptake
Harmsze	Modules	Elementary module	Structuring, organisational and discourse	Implicit (within relations)	No	No	Yes	Open	Presentation of scientific information in electronic articles	Preliminary evaluation
ScholOnto	No	Node, Claim	Cognitive Coherent	Explicit (+ / -)	Explicit (1, 2)	Yes (duplicates)	No	Open	General structuring of coherence and argumentation	Evaluated and widely used
De Waard	Rhetorical Blocks	Rhetorical element	Argumentative	Explicit (within the pairs of relations)	No	No	Yes	Open	Modularization of scientific publications	N/A
SWAN	No	Discourse element, Research statement	Argumentative and Cognitive Coherent	Implicit (within relations)	No	Yes (duplicates)	Yes	Yes (Gene, Protein)	Creation of a knowledge base	Evaluated and widely used
SALT	Rhetorical Blocks	Rhetorical element (Nucleus, Satellite, Claim, Support)	Rhetorical and argumentative	Implicit (within relations)	No	Yes (pointers)	Yes	Open	Structuring of the rhetoric and argumentation in scientific publications	Evaluated and infrequently used

Fig. 1. Comparative overview of the five analyzed discourse representation approaches

Another remaining open question is the set of relations used to connect the elementary discourse knowledge items, as this is the point where the divergence between the existing models is the biggest. Having a closer look at the five sets of relations, we observe two distinct tendencies which can lead to a common denominator. On one side we have a mixture of cognitive coherent and argumentative relations (in the Scholarly Ontologies project, SWAN and Harmsze), while on the other side we have a more linguistic approach materialized in the rhetorical relations used by SALT. Both directions can be used in a complementary fashion. After a refinement of the rhetorical relations, we envision a co-existence of both sets, one modeling the argumentative support of the discourse, while the other capturing the coherence and rationale of the argumentation.

From the properties that relations can carry, we believe that *polarity* should be featured in the unified model, as it is extremely useful both for analysis and visualization of the discourse. The relations' weights are dependent on the extraction mechanism, and therefore should be defined by the corresponding approach and not included in the model, as such discrete quantifiers do not really provide a direct added value for an author / reader. The model also needs to contain the provenance information in addition to the shallow metadata describing the authorship and references.

Finally, the most important “non-functional” element to be considered when designing such an unified model is the adoption from the existing models of the lessons learned with regards to evaluation and uptake. The practical evaluation of the features to be selected for the model should play a crucial role in the overall design. Consequently, the resulting framework needs not only to be elegant and to satisfy all the requirements of a proper formal externalization, but also to be attractive for the average Web user. Contrarily, it will fail to achieve an appropriate community uptake and will remain just an elegant model on paper.

5 Conclusion

In this paper we presented a succinct overview of five of the existing discourse representation models: de Waard's and Harmsze's models, ScholOnto, SWAN and SALT. In addition, we have also made a brief comparative analysis of their main features in terms of organizational structure, types and attributes of the relations between the discourse knowledge items and openness to complementary models or domain knowledge.

Starting from the guidelines proposed in our discussion, we intend to pursue our goal of designing an unified discourse representation model. The next steps will include a series of open discussions with the members of our community and the creation and exposure of a common model, achieved via a shared agreement and understanding.

Acknowledgments

The work presented in this paper has been funded by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (Lion-2).

References

1. Nonaka, I., Takeuchi, H.: *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*. Oxford University Press (1995)
2. Kunz, W., Rittel, H.: Issues as elements of information system. Working paper 131, Institute of Urban and Regional Development, University of California (1970)
3. Sanders, T.J.M., Spooren, W.P.M., Noordman, L.G.M.: Coherence Relations in a Cognitive Theory of Discourse Representation. *Cognitive Linguistics* **4**(2) (1993) 93–133
4. Mann, W.C., Thompson, S.A.: *Rhetorical Structure Theory: A theory of text organization*. Technical Report RS-87-190, Information Science Institute (1987)
5. Hovy, E.: Automated discourse generation using discourse structure relations. *Artificial Intelligence* (63) (1993) 341–385
6. Ciccarese, P., Wu, E., Wong, G., Ocana, M., Kinoshita, J., Ruttenberg, A., Clark, T.: The SWAN biomedical discourse ontology. *J. of Biomedical Informatics* **41**(5) (2008) 739–751
7. Groza, T., Handschuh, S., Möller, K., Decker, S.: SALT - Semantically Annotated L^AT_EX for Scientific Publications. In: *Proceedings of the 4th European Semantic Web Conference (ESWC 2007)*, Innsbruck, Austria (2007)
8. Mancini, C., Shum, S.J.B.: Modelling discourse in contested domains: A semiotic and cognitive framework. *International Journal of Human-Computer Studies* **64**(11) (2006) 1154–1171
9. Harmsze, F.A.P.: *A modular structure for scientific articles in an electronic environment*. PhD thesis, University of Amsterdam (2000)
10. de Waard, A., Tel, G.: The ABCDE format - enabling semantic conference proceeding. In: *Proceedings of 1st Workshop: "SemWiki2006 - From Wiki to Semantics"*, Budva, Montenegro. (2006)
11. Teufel, S., Moens, M.: Summarizing Scientific Articles: Experiments with Relevance and Rhetorical Status. *Computational Linguistics* **28** (2002)
12. Mizuka, Y., Korhonen, A., Mullen, T., Collier, N.: Zone analysis in biology articles as a basis for information extraction. *International Journal of Medical Informatics* **75** (2006) 468–487
13. Lisacek, F., Chichestera, C., Kaplan, A., Sandor, A.: Discovering Paradigm Shift Patterns in Biomedical Abstracts: Application to Neurodegenerative Diseases. In: *Proceedings of the 1st International Symposium on Semantic Mining in Biomedicine (SMBM)*. (2005)
14. Shum, S.J.B.: *Cohere: Towards Web 2.0 Argumentation*. In: *Proceedings of 2nd International Conference on Computational Models of Argument*, IOS Press (2008)
15. Teufel, S., Carletta, J., Moens, M.: An annotation scheme for discourse-level argumentation in research articles. In: *Proc. of the 9th Conf. on European Chapter of the ACL*, Morristown, NJ, USA, ACL (1999) 110–117
16. de Waard, A., Buitelaar, P., Eigner, T.: Identifying the Epistemic Value of Discourse Segments in Biology Texts. In: *Proc. of the 8th Int. Conf. on Computational Semantics (IWCS-8 2009)*. (2009)
17. Groza, T., Möller, K., Handschuh, S., Trif, D., Decker, S.: SALT: Weaving the claim web. In: *Proceedings of the 6th International Semantic Web Conference (ISWC 2007)*, Busan, Korea (2007)
18. Ogden, C.K., Richards, I.A.: *The Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*. Magdalene College, University of Cambridge (1923)