An Implemented, Integrative Approach to Ontology-Based NLP and Interlingua

Marjorie McShane, Sergei Nirenburg and Stephen Beale

Working Paper 06-05

March 8, 2005

Institute for Language and Information Technologies
University of Maryland Baltimore County

Abstract

We describe ways in which the OntoSem text-processing environment integrates two of the realms of interest of this workshop – ontology and interlingua – and takes a different approach to solving the same problems as are tackled by the third – controlled language. We focus on the ways in which recent advances in the three fields have contributed to our thinking about, and the actual development of, the OntoSem environment, and the ways in which we believe that our advances could fundamentally affect the interpretation of realistic goals for NLP.

Introduction

Semantic-rich NLP has been tackled over the years by many systems that wed, with various emphases, work on semantic microtheories, representation terms, interlingua, ontologies and approaches to the central problem of ambiguity resolution. While one can distinguish communities based a primary area of interest, real systems know no such theoretical boundaries. In this paper we discuss a semantic-rich text processing environment that has become known as OntoSem, which incorporates an ontology and related resources aimed specifically at supporting multi-lingual text processing (the original goal of the approach was interlingual machine translation). This OntoSem ontology is linked with lexicons for each language processed, whose semantic structures represent another level of interlingual semantic representation. Among its other goals, OntoSem aims to solve the same problems of ambiguity resolution as controlled vocabularies do, but reaching over the entire vocabulary of a language. Unambiguous, language-neutral semantic representation is expressed in what we call Text-Meaning Representations (TMRs), which are automatically generated representations written in an ontology-grounded metalanguage.

The OntoSem ontology is linked with lexicons for each language processed, whose semantic structures represent another level of interlingual semantic representation. Among its other goals, OntoSem aims to solve the same problems of ambiguity resolution as controlled vocabularies do, but reaching over the entire vocabulary of a language. Unambiguous, language-neutral semantic representation is expressed in what we call Text-Meaning Representations (TMRs), which are automatically generated representations written in an ontology-grounded metalanguage.

Since a comprehensive description of the OntoSem environment lies outside the scope of this paper (see, e.g., Nirenburg and Raskin, forthcoming, and the papers and tutorials at the ILIT website: http://ilit.umbc.edu), we will focus on aspects of OntoSem that tie it into the broader work in the three subfields of interest:

- With respect to **ontology**, we will describe the nature of the OntoSem ontology, our experiments in incorporating external ontologies, the expected direct benefits of the latter and planned enhancements to imported resources.
- With respect to **controlled languages**, we will argue that our TMRs can achieve the same benefits of ambiguity resolution as controlled languages without imposing constraints on the authors of texts.
- With respect to **interlingua**, we will show how the text-meaning representations generated from the OntoSem syntactic-semantic analyzer represents a fine-grained language-neutral semantic representation that can support many multi-lingual applications, including MT.

In short, we will describe a quiet (perhaps too quiet), long-term effort that has been working at namely the type of integration proposed by this workshop.

The OntoSem Ontology

The OntoSem ontology is a language-independent, tangled tree of concepts that represents meanings of objects and events using its own metalanguage. Each concept (there are currently 6000) is described by an average of 16 properties ("features"), selected from the hundreds of properties defined in the ontology. For example, the locally-defined (not inherited) properties for the concept PERFORM-SURGERY include the following (ontological concepts are in small caps; not all fillers are show for reasons of space; the facet 'sem' indicates the typical selectional restriction, 'default' indicates a strongly preferred default, 'inv' indicates the inverse of a property specified elsewhere in the ontology, and 'relaxable-to' indicates a relaxed but still valid interpretation of selectional restrictions):

PERFORM-SURGERY

DEF	VALUE	A type of medical service where an animal is cut open to treat disease or injury.
IS-A	VALUE	TREAT-ILLNESS
SUBCLASSES	VALUE	COLECTOMY
		MESENTERIC-LYMPHADENECTOMY
AGENT	SEM	SURGEON
	RELAXABLE-TO	DOCTOR
INSTRUMENT	INV	SCALPEL
BENEFICIARY	SEM	MEDICAL-PATIENT
DOMAIN-OF	INV	REMEDY-FOR
HAS-EVENT-AS-PART	SEM	CUT

LOCATION	SEM	OPERATING-ROOM
REMEDY-FOR	SEM	ANIMAL-DISEASE

The number of concepts is intentionally restricted so that mappings from lexicons are many-to-one. The concepts are named using quasi-English terms in order to make the resource usable by developers (cf. the MESH/UMLS ontology, discussed below, where the numbered concepts are opaque to human inspection). The OntoSem ontology is geared toward description of the world for the processing – in large part, disambiguation – of language; as such, it is neither a simple hierarchy (like WordNet) nor an attempt to encode all basic world knowledge (like Cyc).

What distinguishes the OntoSem ontologies from all others is the richness of description using properties, which, in conjunction with the ontologically-linked lexicon, supports disambiguation of text, as shown in the next subsection.

Before describing the OntoSem ontology further, let us say a few words about the term ontology. For us, an ontology is language-neutral, property-rich world model, whose use in NLP must be mediated by a lexicon for any given language. Compare this with, for example, Hovy's (1998) definition: "An ontology is set of terms, associated with definitions in natural language (say, English) and, if possible, using formal relations and constraints, about some domain of interest... Under this formulation, a termset, a data dictionary, and even a metadata collection are all types of ontologies. We specifically want it this way because we aim to incorporate information from any of these sources, if appropriate, into the Reference Ontology."

Hovy's highly inclusive definition of ontology, which his group has worked to construct (e.g., in the SENSUS project[1]) has the benefit of avoiding duplication of effort: if an "ontology" of a domain exists, an environment like SENSUS can incorporate it. However, we see three drawbacks to such an omnivorous approach to ontology building.

First, much work must be devoted to the automated merging of ontologies, which is highly error-prone and, due to the size of a resource like SENSUS, will likely not ever be manually corrected. For example, the results of an experiment in merging the top levels of SENSUS and MIKROKOSMOS (the predecessor of OntoSem) were reported as follows: "the combined heuristics extracted 883 suggestions for validation (= 2.72% of the total number of pairs, or 13% of the portion of SENSUS under consideration). Of these, 244 (= 27.6%) were correct, 383 (= 43.4%) were incorrect, and 256 (= 30.0%) were nearly correct" (Hovy 1998). Although the alignment task is undoubtedly very difficult, it is still not clear what can be done with results that are only about 1/3 correct. We see a parallel here with the amount of time the community has spent trying to learn to use WordNet rather than building a resource that is intended for NLP (see Nirenburg et al 2004b). Low-quality (that is, errorful) resources will yield low-quality application results; and no amount of broad coverage will help to raise the quality of output if the coverage is prone to error and not sufficiently deep. We appreciate the need for short-term progress. We doubt that the results of work oriented at short-term goals will be in practical terms useful for the work that strives to attain higher quality levels in text understanding and its applications.

Second, each resource presents its own, different, problems. For example, SENSUS is "a rearrangement and extension of WordNet... retaxonomized under the Penman Upper Model..." (Hovy 1998). Fellbaum (1998, 1999a,b), a developer of WordNet, has reported many insufficiencies of the resource for use in NLP, and we have identified some more (Nirenburg et al 2004b), the most devastating being the inability to support disambiguation. Such ambiguity is carried over into SENSUS since, as in WordNet, each node is an English word. Therefore, where one gains in not having to reinvent a word net, one loses in that a word net is far from an ideal resource for NLP.

Third, it is not at all clear that a bigger ontology is a better ontology. The obvious parallel is with sense splitting versus sense bunching on the lexical level. Whereas it is customary to assume that a larger resource implies better coverage, this assumption is known to be flawed – one of the reasons attempts to use MRDs in the 1990s have been all but abandoned (see, e.g., Ide and Véronis 1993). In addition, lexicalizing metaphorical usage, as is done in WordNet, not only doesn't help in automatic disambiguation, it actually ambiguates texts that would otherwise be straightforward (e.g., listing "heart" in the meaning "beloved person").

The OntoSem ontology, by contrast, is handcrafted and property rich. It represents *a* (not the only possible) view of the world that we have found useful for NLP. We stick by this core, building it as necessary but not assuming (as is typical and understandable for, say, Semantic Web developers) that all ontologies are useful by virtue of their very existence (a great simplification of that line of study, but a reasonable snapshot of the basic approach). However, this does not mean that we are uninterested in externally developed resources – we are, but in the circumscribed ways we now discuss.

We are pursuing full-scale importation of external ontologies but only in specialized domains, not as competing world views for our core. Our current experiment involves culling from the MESH and UMLS medical resources that information that we consider useful for expanding our ontology and English lexicon into the medical domain.[2] We have extracted a hierarchical tree of

[1] See http://www.isi.edu/naturallanguage/projects/ONTOLOGIES.html

[2] See http://www.nlm.nih.gov/mesh/MBrowser.html and http://www.nlm.nih.gov/research/umls/UMLSDOC.HTML#s01.

about 250,000 concepts, with an average of 4 lexical items (representing different terminological schemes) linked to each. Some, albeit limited, properties are defined for concepts, which we are currently exploring.

The benefit of importing specialized ontologies and lexicons is clear: a huge savings in effort for gaining specialized-domain knowledge. There is, however, an associated wish list: i.e., we would like to have the resources to handcraft, or the ability to machine-learn, properties for these concepts that not only distinguish them, but also better interconnect them with each other and with our base ontology. This would permit imported ontologies to be on a par of quality and grain-size with our base ontology.

A question: what do we gain by importing an "underspecified" (i.e., property-poor) ontology into OntoSem? When the various branches of the imported ontology are connected to the appropriate leaves in the OntoSem ontology (which is done manually), everything that is known about the anchor OntoSem concept is propagated via inheritance to all the concepts in that subtree. So, e.g., an imported subtree of surgical procedures will be linked to the OntoSem concept PERFORM-SURGERY, and everything the system knows about the AGENT, INSTRUMENT, BENEFICIARY, LOCATION, etc., of PERFORM-SURGERY will be propagated to the new descendants. Of course, we'd prefer to have a surgical procedure like "appendectomy" supplied with information like:

```
APPENDECTOMY
    HAS-EVENT-AS-PART     REMOVE
                                THEME   APPENDIX
```

or, more simply but still informative:

```
APPENDECTOMY
    THEME                 APPENDIX
```

but just knowing that an appendectomy is a surgical procedure already tells us much of what we need to know about it for the purpose of making an unambiguous text-meaning representation.

Another way in which we are using external resources is to inform manual knowledge acquisition. For example, we have been using WordNet – as well as the paper resource *Word Menu* and dictionary.com – to drive synonym- and hyponym-based lexicon (and associated ontology) acquisition. Such resources are valuable raw material that is massaged by knowledge acquirers to reflect what we believe will be most useful for our practical goals. While massaging a resource is obviously more costly than directly importing it, the process is not prohibitively expensive (at least outside of specialized domains, like the medical one), and it does the job once and for all, rather than postponing error correction until some unspecified future stage of system development.

The OntoSem Lexicon: Another Level of Interlingua

The OntoSem environment uses ontologically-linked lexicons (for each language processed) to mediate between texts and the ontology. Each lexical sense contains (among other information) syntactic and semantic zones, linked through special variables, as well as procedural-semantic attachments that we call "meaning procedures", which compute context-specific meanings on the fly. The current general-domain lexicon contains 18,000 entries, including both word-level and phrase-level entities. We also have an onomasticon of several hundred thousand entries.

A basic verbal lexicon entry in OntoSem looks as follows, in presentation format:

```
watch
watch-v1
  synonyms "observe"
  anno
     definition  "to observe, look at"
     example "He's watching the competition."
syn-struc
    subject        $var1     cat n
     v             $var0     cat v
    directobject   $var2     cat n
sem-struc
   VOLUNTARY-VISUAL-EVENT
       agent       ^$var1
       theme       ^$var2
```

The syntactic structure (syn-struc) says that this is a transitive sense of *watch*. The semantic structure (sem-struc) says that a VOLUNTARY-VISUAL-EVENT – which is a concept in our ontology – must be instantiated in the text-meaning representation (TMR), with its agent and theme being realized by the subject and direct object, respectively (as shown by the linked variables; ^ is read 'the meaning of').

Apart from mapping directly to an ontological concept, there are many other ways to express meaning in OntoSem. For example, one can map to an ontological concept with modified property values: e.g.,

- **Zionist** is described as a POLITICAL-ROLE that is the AGENT-OF a SUPPORT event whose THEME is Israel.
- **asphalt (v.)** is described as a COVER event whose INSTRUMENT is ASPHALT.
- **recall** (**v.** as in *they recalled the high chairs*) is described as a RETURN-OBJECT event that is CAUSED-BY a FOR-PROFIT-CORPORATION and whose THEME is ARTIFACT, INGESTIBLE or MATERIAL.

There are also a number of fully or partially non-ontological ways of describing meaning, like the use of parametric values of mood or aspect. For example, the auxiliary *might* as in *He might come over* is described using the modality 'epistemic', which deals with the truth value of a statement:

```
  syn-struc
      subject      $var1    cat n
      v            $var0    cat v
      inf-cl       $var2    cat v
  sem-struc
    ^$var2
      epistemic   .5
      agent       ^$var1
  meaning-procedure
    fix-case-role (value ^$var1) (value ^$var2) [3]
```

Another set of extra-ontological semantic descriptors is used for time expressions, as shown by the example of *yesterday* below.

```
syn-struc
   root     $var1       cat v
   mods     root $var0  cat adv
            type        pre-verb-post-clause
sem-struc
  ^$var1
      time
         combine-time
            (find-anchor-time) (day 1) before
```

As shown in the examples of *might* and *yesterday*, calls to procedural semantic routines are used widely in OntoSem lexical description. This reflects the fact that many aspects of meaning cannot be statically described but, rather, must be computed. An advantage of developing lexical resources within a processing environment is being able to assign responsibility for portions of semantic composition to resources best suited for them.

In addition to the means of lexical expression described above, OntoSem lexicon entries can include entities of any degree of complexity, including phrasals of any profile, as reported in McShane et al. (ms.).

What must be emphasized is how language neutral – and therefore portable across languages – the semantic descriptions in the OntoSem lexicon are. Whereas it is typical to assume that lexicons are language-specific whereas ontologies are language-independent, most aspects of OntoSem sem-strucs are language-independent, apart from the linking of specific variables to their counterparts in the syn-struc. Stated differently, if we consider sem-strucs – no matter what lexicon they originate from – to be building blocks of the representation of *word meaning* (as opposed to concept meaning, as is done in the ontology), then the job of writing a lexicon for L2 based on the lexicon for L1 is in large part limited to a) providing an L2 translation for the head word(s), b) making any necessary syn-struc adjustments and c) checking/modifying the linking among variables in the syn- and sem-strucs (McShane et al. 2004). This conception of cross-linguistic lexicon development derives in large part from the Principle of Practical Effability (Nirenburg and Raskin 2004), which states that what can be expressed in one language can *somehow* be expressed in all other languages, be it by a word, a phrase, etc.

Apart from this theoretical justification for conceptualizing the sem-strucs as building blocks for lexical representation, there are two practical rationales: supporting consistency of meaning representation across languages and using acquirer time most efficiently in large-scale lexical acquisition.

As regards consistency, the potential for paraphrase must be considered when building multi-lingual resources. For instance, 'weapons of mass destruction' can be described as the union of CHEMICAL-WEAPON and BIOLOGICAL-WEAPON, or it can be described as WEAPON with the ability to KILL > 10,000 HUMANs (the actual number recorded will be treated by the analyzer in a fuzzy fashion; however, it would be less than ideal for a lexicon for L2 to record 10,000 while a lexicon for L3 recorded 25,000). While both representations are valid, it is desirable to use the same one in all languages covered. In addition, the decision of how to describe a notion – whether by ontologizing it, describing it using extra-ontological means, describing it using an existing concept with additional properties and values defined – is often a judgment call. It would not be desirable for the acquirer of German to map the word *Schimmel* 'white horse' to the concept HORSE with the lexical restriction COLOR: WHITE, while the acquirer of some other language that also has a word for 'white horse' introduced an ontological concept specifically for this entity. Again, while both representations are valid and semantically equivalent, the general tendency should be to strive toward uniformity where possible.

As concerns acquirer time, composing sem-strucs is, by far, the most time- and effort-intensive aspect of writing OntoSem lexicon entries. This derives from the wealth of expressive means; the fact that microtheories of time, reference, etc., are developed during lexicon acquisition (recall that our environment is fully integrated with processors); and the fact that ontology development occurs hand-in-hand with lexicon development. Therefore, work on the first lexicon entry that describes a word sense – regardless of the language of origin – takes much more time than editing a word sense for a new language. Moreover, although in the worst case some editing of entries is necessary for L2, L3, etc., in most cases no such editing is needed.

[3] This meaning procedure reassigns a case-role if the listed AGENT case-role is inappropriate considering the meaning of $var1 and/or $var2: e.g., in *the truck might come,* truck is a THEME of a MOTION-EVENT, not an AGENT, and in *I might get sick,* I am an EXPERIENCER of a DISEASE event, not an AGENT of it.

In its focus on unambiguous text-meaning representation and cross-linguistic manipulation, the OntoSem environment is reminiscent of projects in the field of controlled languages. However, the goal in OntoSem is more lofty (to achieve disambiguation without the need for controlled input vocabulary) and the research program, accordingly, more long-term. Consider, for example, the following lexial senses of 'see', which are listed in abbreviated form for reasons of space: 'tr.' 'intr.' and 'bitr.' indicate transitive, intransitive, and bitransitive syntactic structures, respectively; the use of bold, italics and underlining show the correspondence of syntactic entities and semantic roles using real examples, for the sake of clarity.

see-v1 (tr.) *He* saw **her new car**
VOLUNTARY-VISUAL-EVENT
AGENT
THEME (PHYSICAL-EVENT, PHYSICAL-OBJECT, SOCIAL-EVENT)

see-v2 (intr.) I can't help you. – **I** see.
UNDERSTAND
EXPERIENCER (HUMAN)

see-v3 (tr.) For further information, see **chapter 8**.
READ
AGENT (HUMAN)
THEME (TEXT-UNIT)

see-v4 (tr.) *Grandma* saw **her doctor** yesterday.
CONSULT
AGENT (HUMAN)
THEME (MEDICAL-ROLE, LEGAL-ROLE)

see-v5 (bitr.) <u>He</u> saw *her* **to her car**.
ESCORT
<u>AGENT (HUMAN)</u>
THEME (HUMAN)
DESTINATION

Disambiguation of these senses is carried out by the OntoSem analyzer, using the combination of syntactic and semantic diagnostics (see Beale et al. 2003). Thus, if the theme of a transitive structure semantically represents a TEXT-UNIT (as Chapter 8 will, once it is analyzed semantically), sense 3 will be selected (the implied agent of the imperative is handled by special rules); similarly, if the syntactic structure is bitransitive with a HUMAN theme and a destination, then the meaning ESCORT (v5) will be selected. While disambiguation does not succeed in 100% of cases, our results have been quite good (see Nirenburg et al. 2004a for evaluation of OntoSem).

Ontology + Lexicon (+ Analyzer) → TMR: Three Levels of Interlingua

As we have already shown, the OntoSem ontology is fully language-independent, and the related lexicons are largely language-independent, apart from some language-specific syntactic structures, some potentially idiosyncratic syntax-semantics linkings and, of course, the actual strings that realize lexical items in different languages. But perhaps the most important aspect of the OntoSem interlingua are the text-meaning representations (TMRs) that are both the results of processing raw text and the input to text generation (for MT), summarization, QA, etc.

Below is a simple TMR from a recently processed text, in presentation format (simple because most of the sentences we process are much longer). It reflects the meaning of the sentence **He asked the UN to authorize the war**. It is written in the TMR language, which is a metalanguage for representing text meaning that is compatible with the metalanguages used in the lexicon, ontology and fact repository (the latter being a database of real-world facts extracted from texts).

REQUEST-ACTION-69	
AGENT	HUMAN-72
THEME	ACCEPT-70
BENEFICIARY	ORGANIZATION-71
SOURCE-ROOT-WORD	ask
TIME	(< (FIND-ANCHOR-TIME))
ACCEPT-70	
THEME	WAR-73
THEME-OF	REQUEST-ACTION-69
SOURCE-ROOT-WORD	authorize
ORGANIZATION-71	
HAS-NAME	UNITED-NATIONS
BENEFICIARY-OF	REQUEST-ACTION-69
SOURCE-ROOT-WORD	UN
HUMAN-72	
HAS-NAME	COLIN POWELL
AGENT-OF	REQUEST-ACTION-69
SOURCE-ROOT-WORD	he; *ref. resolution done*
WAR-73	
THEME-OF	ACCEPT-70
SOURCE-ROOT-WORD	war

This says that there is a REQUEST-ACTION event whose AGENT is HUMAN-72 (Colin Powell), whose BENEFICIARY is ORGANIZATION-71 (United Nations) and whose THEME is ACCEPT. The ACCEPT event, in turn, has a THEME of WAR-73. Note that the concept ACCEPT is *not* the same as the English word *accept*: its human-oriented definition in the ontology is "To agree to carry out an action, fulfill a request, etc", which fits well here. So, TMRs are unambiguous representations of the meaning of a text, grounded in a language-independent world model (the ontology). They represent, to our knowledge, the most advanced level of real-scale semantic representation of texts achieved by any system so far.

The Big Picture

As this brief overview of OntoSem has shown, we as developers of the system certainly believe that the three realms under investigation in this workshop – interlingua, ontologies and controlled languages – are most distinctly related and their relations must be pursued.

Within each field of study, feasibility has been of central concern. Research in the Semantic Web proceeds from the practical notion that the Web *is*, that it is too big to be hand manipulated, and that its wide array of information must be managed the best we can keeping size, user profiles, etc., in mind (see Berners-Lee et al. 2001). Research in controlled languages tackles the major hurdle of ambiguity in the short-term by restricting the scope of the problem by imposing limitations on language users in domains where such limitations can be enforced. And research in interlingua attempts to exploit the similarity of what can be expressed in natural languages in order to offset the need to build NLP systems for every imaginable language pair.

In OntoSem, issues of practicality are no less pressing. Building the high-quality resources our processors use is expensive and time-consuming – but, as we have argued elsewhere (e.g., Nirenburg et al. 2004b), no more expensive or time-consuming than trying to leverage resources that weren't intended to support NLP. To alleviate the expense of hand-acquiring knowledge, we are pursuing various avenues, like manually exploiting certain high-cost resources (like WordNet), selectively incorporating others (like MESH/UMLS) with various degrees of hand massaging, and learning to use machine learning to our advantage in ontological and lexical acquisition (a topic that strays from the focus of this paper but is well under way in our research program).

We believe that in our research program we have achieved the critical mass of high-quality knowledge resources and robust processors to make the goal of automatically producing unambiguous TMRs from unrestricted input text – which is generally dismissed as impossible – a reality in the near future (we cannot yet claim the production-level results of the controlled language community). The availability of automatically generated TMRs alleviates many of the problems in applications like MT and QA. The increasing availability of resources like WordNet, MESH and many other such can drive this process when used in closely monitored ways.

References

Beale, Stephen, Sergei Nirenburg and Marjorie McShane. 2003. Just-in-Time grammar. Proceedings 2003 International Multiconference in Computer Science and Computer Engineering. Las Vegas, Nevada.

Berners-Lee, T, J. Hendler and O. Lassila. 2001. The Semantic Web. *Scientific American*, May 17, 2001

Fellbaum, C. 1999a. A semantic network of English: The mother of all wordnets. Computers and the Humanities 32: 209-220.

Fellbaum, C. 1999b. Verb semantics via conceptual and lexical relations. In E. Viegas (Ed.), Breadth and Depth of the Lexicon (pp. 247-262). Dordrecht, Holland: Kluwer Academic Publishers.

Glazier, S. 1998. *Random House Webster's Word Menu*. Random House.

Hovy, E. H. 1998. Combining and Standardizing Large-Scale, Practical Ontologies for Machine Translation and Other Uses. Proceedings of the International Conference on Language Resources and Evaluation (LREC). Granada, Spain.

Ide, N. and J. Véronis. 1993. Extracting knowledge bases from machine-readable dictionaries: Have we wasted our time? In Proceedings of *KB&KS'93 Workshop,*Tokyo, 257-266.

McShane, M., M. Zabludowski, S. Nirenburg, and S. Beale. (2004). OntoSem and SIMPLE: Two Multi-Lingual World Views. Proceedings of ACL-2004 Workshop on Text Meaning and Interpretation. Barcelona, Spain.

McShane, M., S. Nirenburg, and S. Beale. (Ms.). The Description and Processing of Multiword Expressions in OntoSem.

Nirenburg, S., S. Beale, and M. McShane. (2004a). Evaluating the Performance of the OntoSem Semantic Analyzer. Forthcoming in Proceedings of the ACL Workshop on Text Meaning Representation.

Nirenburg, S., M. McShane, and S. Beale. (2004b). The Rationale for Building Resources Expressly for NLP. Proceedings of LREC-2004.

Nirenburg, S. and V. Raskin. 2004. *Ontological Semantics*. MIT Press.