Overcoming Linguistic Barriers to the Multilingual Semantic Web

Graeme Hirst

Abstract I analyze Berners-Lee, Hendler, and Lassila's description of the Semantic Web, discussing what it implies for a Multilingual Semantic Web and the barriers that the nature of language itself puts in the way of that vision. Issues raised include the mismatch between natural language lexicons and hierarchical ontologies, the limitations of a purely writer-centered view of meaning, and the benefits of a reader-centered view. I then discuss how we can start to overcome these barriers by taking a different view of the problem and considering distributional models of semantics in place of purely symbolic models.

Key Words Distributional semantics • Near-synonymy • Ontologies • Reader-centered view of meaning • Semantic Web • Writer-centered view of meaning

1 Introduction

> The Semantic Web ... in which information is given well-defined meaning, better enabling computers and people to work in cooperation. — Berners-Lee et al. (2001, p. 37)[1]

Sometime between the publication of the original paper with this description of the Semantic Web and Berners-Lee's (2009) "Linked Data" talk at TED, the vision of the Semantic Web contracted considerably. Originally, the vision was about "information"; now it is only about data. The difference is fundamental. Data, even if it is strings of natural language, has an inherent semantic structure and a stipulated interpretation, even if that too is a label in natural language. Other kinds of information, however, are semi-structured or unstructured and may come with no interpretation imposed. In particular, textual information gains an interpretation only in context and only for a specific reader or community of readers (Fish 1980).

[1]I will refer to these authors, and metonymously to this paper, as *BLHL*.

G. Hirst (✉)
Department of Computer Science, University of Toronto, Toronto, ON, Canada M5S 3G4
e-mail: gh@cs.toronto.edu

P. Buitelaar, P. Cimiano (eds.), *Towards the Multilingual Semantic Web*,
DOI 10.1007/978-3-662-43585-4_1

I do not mean to criticize the idea of restricting Semantic Web efforts to data *pro tem*. It is still an extremely challenging problem, especially in its multilingual form (Gracia et al. 2012, this volume *passim*), and the results will still be of enormous utility. At the same time, however, we need to keep sight of the broader goal that BLHL's vision implies in order to make sure that our efforts to solve the preliminary problem are not just climbing trees to reach the moon. In this chapter, I will perform a hermeneutical analysis of BLHL's description, with discussion of what it implies for the Multilingual Semantic Web and the barriers that the nature of language itself puts in the way of that vision. I will then discuss how we can start to overcome these barriers.

I assume in this chapter the standard received notion of the Multilingual Semantic Web as one in which web pages contain (inter alia) natural language text but are also marked up with semantic annotations in a logical representation that enables inferences to be made, that is independent of any particular natural language, and that draws on shared ontologies that are also language-independent. And consequent upon that, the Multilingual Semantic Web, in response to users' queries and searches, expressed in a natural language or by other means, is able to bring together multiple pages in multiple languages, matching the query to semantic annotations, drawing inferences as necessary, and presenting the results in whatever language the user wants, translating from one language to another as necessary.

2 Well-Defined Meaning and Multilinguality

In BLHL's vision, "information is given well-defined meaning," implying paradoxically that information did not have well-defined meaning already. Of course, the phrase "well-defined meaning" lacks well-defined meaning, but BLHL are not really suggesting that information on the non-Semantic Web is meaningless; rather what they want is *precision* and the *absence of ambiguity* in the semantic layer. In the case of information expressed linguistically, this implies semantic interpretation into a symbolic knowledge representation language of the kind that they talk about elsewhere in their paper. Developing such representations was a goal that exercised, and ultimately defeated, research in artificial intelligence and natural language understanding from the 1970s through to the mid-1990s (Hirst 2013) (see Sect. 5) and which the Semantic Web has made once more a topic of research (e.g., Cimiano et al. 2014).

One of the barriers that this earlier work ran into was the fact that traditional symbolic knowledge representations proved to be poor representations for linguistic meaning and hierarchical ontologies proved to be poor representations for the lexicon of a language (Hirst 2009a).[2] Models such as LexInfo and *lemon*

[2] Wilks (2009), echoed by Borin (2012), suggests that, *a fortiori*, "ontologies" as presently constructed are nothing more than substandard lexicons disguised as something different.

(Cimiano et al. 2011; McCrae et al. 2012) attempt to associate multilingual lexical and syntactic information with ontologies, but they necessarily retain the idea that "the sense inventory is provided by a given domain ontology" (Cimiano et al. 2011, fn 9), under the assumption that the domain of a text, and hence the requisite unique ontology, is known *a priori* or can be confidently identified prior to semantic analysis. In practice, however, this leads to an inflexible and limiting view of word senses. For example, languages tend to have many groups of near-synonyms that form clusters of related and overlapping meanings that do not admit a hierarchical differentiation (Edmonds and Hirst 2002). And quite apart from lexical issues, any system for fully representing linguistically expressed information must itself have the expressive power of natural language, which is far greater than the first-order and near-first-order representations that are presently used; but the higher-order and intensional representations required for this degree of expressiveness (Montague 1974) are computationally infeasible (Friedman et al. 1978).

All these problems are compounded when we add multilinguality as an element. For example, different languages will often present a different and mutually incompatible set of word senses, as each language lexicalizes somewhat different categorizations or perspectives of the world and each language has *lexical gaps* relative both to other languages and to the categories of a complete ontology (Hirst 2009a, pp. 278–279). The consequence of these incompatibilities for the Multilingual Semantic Web is that the utility of ontologies for interpreting linguistic information is thereby limited, and so, conversely, is the ability of lexicons to express ontological concepts. This leads to practical limitations on models of lexicons for ontologies, such as McCrae et al.'s (2012) *lemon* model, that put an emphasis on *interchangeability*—the idea that one ontology can have many different lexicons, for example, for different languages or dialects. This wrongly assumes that *translation-equivalent words* have identical meanings. In fact, it is rare even for words that are regarded as translation equivalents to be completely identical in sense, and such cases are limited mostly to cross-language borrowings and monosemous technical terms in highly structured domains (Adamska-Sałaciak 2013). For example, the sport of soccer, which Cimiano et al. (2014) use as a domain to exemplify an ontology with interchangeable lexicons, is sufficiently technical and well-structured for the approach to succeed; so are the deliberately very narrow domains considered by Embley et al. (this volume). But interchangeability might fail even in ontologies for well-structured domains (cf. Léon-Araúz and Faber, this volume). For example, regarding the domain of university administrative structures, Schogt (1988, p. 97) writes: "When I want to talk about aspects of the intricate administrative system of the University of Toronto to Dutch academics it is very difficult to use Dutch because there are no Dutch terms that correspond to those used in Toronto, the Dutch set-up not sharing the functions and divisions that characterize the Toronto system."

More usually, translation-equivalent words are merely cross-lingual near-synonyms (Hirst 2009a, p. 279). For example, in the concept space of differently sized areas of trees, the division between the French *bois* and *forêt* occurs at a "larger" point than the division between the German *Holz* and *Wald*

(Hjelmslev 1961; Schogt 1976, 1988). Similarly, English, German, French, and Japanese all have a large vocabulary for different kinds of mistakes and errors, but they each divide up the concept space quite differently. For example, the Japanese words that translate the English words *mistake* or *error* include *machigai*, *ayamari*, and *ayamachi*; Fujiwara et al. (1985) note:

> *Machigai* implies a straying from a proper course or the target, and suggests that the results are not right. *Ayamari* describes wrong results objectively. Focus of attention is given solely to the results; concerns, worries, or inadvertence in the course of action are not taken into consideration as in *machigai*. *Ayamachi* implies serious wrongdoing or crime. Also, it is used for accidental faults. *Ayamachi* is concerned with whether the results are good or bad, based on moral judgement, while *ayamari* is concerned with whether the results are right or wrong.[3]

To translate the same two English words *mistake* and *error* to German, Farrell (1977, p. 220) notes that even though *error* "expresses a more severe criticism than *mistake*", both are covered by *Fehler*, except that *Irrtum* should be used if the mistake is a misunderstanding or other mental error and *Mangel* if the mistake is a "deficiency [or] absence" rather than a "positive fault or flaw" or if it is a visible aesthetic flaw.

These kinds of translation misalignments are common across languages. However, in the *lemon* model, we cannot, for example, just have a concept in our ontology for a smallish area of trees, which *bois* and *Holz* map to, and one for a bigger area, which *forêt* and *Wald* map to. Rather, to properly represent the meanings of these words, we must have four separate language-dependent concepts in our ontology. (*lemon* allows language-dependent concepts to be defined for use within a specific lexicon, provided, of course, that the new concept is expressible in terms of the existing external ontology (Cimiano et al. 2014).) Additional languages complicate the picture further; for example, Dutch gives a spectrum of three words, *hout*, *bos*, and *woud* (Henry Schogt, p.c.). A language-independent ontological representation of the different kinds of errors that are lexically reified by various languages, a small sample of which was shown above, would be even more complex. Of course, an ontology may be "localized" to a particular language, as posited by Gracia et al. (2012), but cross-lingual mappings between localized ontologies will be very difficult in practice; the example given by Gracia et al. covers only one easy case where a term in one language neatly subsumes two in another (English *river*, French *fleuve* and *rivière*).

Edmonds and Hirst (2002) have proposed that instead of thus making the ontology ever more fine-grained as additional languages are taken into account, only relatively coarse-grained ontological information should be used in the lexicon, along with explicit differentiating information for nonhierarchically distinguished near-synonyms, both within and across languages—much as we saw in the examples above from Fujiwara et al. and Farrell, albeit in a formal representation. Drawing on this model, Inkpen and Hirst (2006) used the explicit differentiating information

[3]Thanks to Kazuko Nakajima for the translation of this text from Japanese.

in conventional dictionaries and dictionaries of near-synonym explication to develop knowledge bases of lexical differentiation for English and (minimally) for French. Inkpen and Hirst demonstrated that using this knowledge of lexical differences improved the quality of lexical choice in a (toy) translation system, using aligned French–English sentence pairs from the proceedings of the Canadian Parliament as test data. Nonetheless, differentiating information on nonhierarchically distinguished near-synonyms, within or across languages, might need to be used in inferences. A Multilingual Semantic Web cannot rely on only an ontology as an interlingual representation or as a nonlinguistic representation for inference; there is, in practice, no clean separation between the conceptual and the linguistic.

3 Given Meaning by Whom?

In BLHL's vision, "information is given well-defined meaning"—but by whom? BLHL's answer was clear: it would be done by the person who provides the information. "Ordinary users will compose Semantic Web pages and add new definitions and rules using off-the-shelf software that will assist with semantic markup" (BLHL, p. 36). That is, semantic markup—and even the creation of new ontological definitions and rules—is assumed to occur at page-creation time, either automatically or, more usually, semi-automatically with the assistance of the author, who is an "ordinary user"—the writer of a blog, perhaps. Hence, in this view a Semantic Web page has a single, fixed, semantic representation that (presumably) reflects its author's view of what he or she wants or expects readers of the page to get from it. The markup is created in the context of the author's personal and linguistic worldview.

This is a *writer-centered view of meaning*. It assumes that the context, background knowledge, and agenda that any potential user or reader of the page will draw on in understanding its content are the same as those of the author and that therefore the meaning that the user will take from the page is the same as the meaning that its creator put in. This is so both in the case that the user is a human looking at natural language text and in the case that the user is software looking at the semantic markup. It is a version of the *conduit metaphor* of communication (Reddy 1979), in which text (or markup) is viewed as a container into which meaning is stuffed and sent to a receiver who removes the meaning from the container and in doing so comprehends the text and thereby completes the communication. This view may also be thought of as *intention-centered*, in that, barring mistakes and accidents, the meaning received is the meaning that the author intended to convey.

Many potential uses of the Semantic Web fit naturally into the paradigm of markup for writer-based meaning and an intention-centered view. These uses are typically some kind of *intelligence gathering*, in the most general sense of that term—understanding what someone else is thinking, saying, or doing. That is, the user's question, looking at some text, is "What are they trying to tell me?"

(Hirst 2007, 2008). Tasks that fit this paradigm, in addition to simple searches for objectively factual information, include sentiment analysis and classification, opinion extraction, and ideological analysis of texts—for example, finding a well-reviewed hotel in a particular city. In each of these tasks, determining a writer's intent is the explicit goal, or part of it, and the markup will help to do this.

Future methods of automatic translation of Semantic Web pages also fall under this paradigm. The goal of translation is to reproduce the author's intent as well as possible in the target language. Translation systems will be able to use both the original natural language text and the author's markup in order to produce a translation that is more accurate and more faithful to the author's intent than a system relying on the text alone could produce.

However, this writer- and intention-centered view is too constraining and restrictive for fully effective use of the Semantic Web—in fact, for many of the primary uses of the Semantic Web. Consider, for example, the limitations that this view puts on search. For a search to usefully take domain circumscriptions and shared ontologies into account, the user must be thinking and searching in the same terms as those of the author of the information that the user wishes to find. If there is a conceptual mismatch, then the information sought might not be found at all—an outcome no better than a simple string-matching search with unfortunately chosen terms.[4] And this leads to my next point.

4 Work Together for Whose Benefit?

In BLHL's vision, the Semantic Web will "better [enable] computers and people to work in cooperation [with each other]." But for whose benefit is this? The Semantic Web vision rightly emphasizes the benefit of the *information seeker*, whose task will be made easier and who will be given a greater chance of success. The benefit to the *information provider*, who wants to bring their information to the notice of the world for commercial, administrative, or other purposes, is secondary and often indirect.

And this is why a strictly writer-based view of meaning is inappropriate for the Semantic Web. Much of the potential value of querying the Semantic Web is that the system may act on behalf of the user, finding relevance in, or connections between,

[4]For example, contemporary researchers in biodiversity have trouble searching the legacy literature in the field because diachronic changes both in the terminology and in the conceptual understanding of the domain result in there being no shared ontologies. "Even competent and well-intentioned researchers often have difficulties searching this literature. Simple Google-style keyword searches are frequently insufficient, because in this literature, more so perhaps than most other fields of science, related concepts are often described or explained in different terms, or in completely different conceptual frameworks, from those of contemporary research. As a result, interesting and beneficial relations with legacy publications, or even with whole literatures, may remain hidden to term-based methods" (Hirst et al. 2013).

interpretations of texts that go beyond anything that the original authors of those texts intended. For example, if the user wants to find, say, evidence that society is too tolerant of intoxicated drivers or evidence that the government is doing a poor job or evidence that the Philippines has the technical resources to commence a nuclear-weapons program, then a relevant text need not contain any particular set of words nor anything that could be regarded as a literal assertion about the question (although it might), and the writer of a relevant text need not have had any intent that it provide such evidence (Hirst 2007, p. 275).

But for a Semantic Web system to find situations in which a document *unintentionally* answers an information seeker's query, it must embody also a *reader-centered view of meaning*. It must be able to ask, on behalf of the user, "What does this text mean to me?" (Hirst 2007, 2008). In its most general form, this is a postmodern view of text, in which the interpretation of each reader, or each community of like-minded readers, may be different (Fish 1980). Here, however, we need only a more limited view: that the system understand the user's goal or purpose in their search and, ideally, the user's viewpoint, beliefs, or ideology and "anything else known about the user" (Hirst 2007, p. 275). That is, a *user model* is available to the system, and, moreover, an agent local to the user's search interface has possibly inferred (or been explicitly told) the broader context or purpose of the user's current activity. The elements of the user model might, in turn, be partially derived or inferred from the system's observation of the user's prior reading and prior searches, in addition to feedback and possibly explicit training from the user. It would start as a generic model and then adapt and accommodate itself to the individual user, becoming more precise and refined (Hirst 2009b). In particular, the user model might include aspects of the user's beliefs and values and their reflections in ontology and lexis—for example, which shared ontologies the user accepts and which ones they reject. These factors may then be used in the search to answer the user's query, perhaps becoming part of the query itself and being used in matching and inference processes to interpret Semantic Web pages.

Consider, for example, a user who wants to know whether they should spend their time and money on a certain movie. A writer-centered Semantic Web would require them to ask a *proxy question* such as "Did other people like this movie?", whereas a reader-centered Semantic Web would allow them to ask their real question, "Will *I* like this movie?". If the system knows, from its model of the user, that they prefer quiet, intelligent movies, then a disgruntled review criticizing the movie for its lack of action could be interpreted as a positive answer to the question. More generally, a reader-centered perspective is particularly useful for abstract, ideological, wide-ranging, or unusual questions and for tasks such as nonfactoid question-answering and query-oriented multi-document summarization where interpretation is an essential part of the task.

Of course, as the movie example above suggests, it may still, in the end, be the writer's annotation to which a reader-centered matching process will be applied. However, the writer's annotation need no longer be the only annotation considered. Whenever a user's query matches a page, the retrieval software may add an annotation to that page with the reader-centered interpretation and inferences that are produced and the reader characteristics upon which they are based. This will facilitate future matching by similar readers with similar queries. Thus, in time a Semantic Web page might bear many different annotations reflecting many different interpretations, not merely that of the writer.[5] In particular, for the Multilingual Semantic Web, these annotations may include translations and glosses that future processes may use.

None of this is to say that the writer-centered view isn't valuable too; as we noted earlier, many intelligence-gathering tasks fit that paradigm. The ideal Semantic Web would embody both views. And the ontolexical resources, markup, and inference mechanisms of a writer-centered Semantic Web are a prerequisite for the additional mechanisms of a reader-centered view.

5 Overcoming Linguistic (and Representational) Barriers

The discussion above gives us a starting point for thinking about what our next steps should be toward a monolingual or Multilingual Semantic Web that includes textual information. First, it implies that we must, in some ways, lower our expectations. We must give up, at least *pro tem*, the goal of creating a Semantic Web that relies on high-quality knowledge-based semantic interpretation and translation or understanding across languages. We must accept that any semantic representation of text will be only partial and will be concentrated on facets of the text for which a first-order or near-first-order representation can be constructed and for which some relatively language-independent ontological grounding has been defined. Hence, the semantic representation of a text may be incomplete, patchy, and heterogeneous, with different levels of analysis in different places (Hirst and Ryan 1992). We must also accept that the Semantic Web will be limited, at least in the initial stages, to a static, writer-centered view of meaning.

However, we should *not* take the view that the Semantic Web will remain "incomplete" until BLHL's vision is realized. Rather, we should say that at each step along the way it will on the one hand be a useful artifact and on the other hand will remain "imperfect." The difference is that an *incomplete* Semantic Web would be missing certain features or abilities but would be fully realized in other respects; the underlying metaphor is one of piece-by-piece construction from components that are each already individually complete and perfect at the time that they are added, and the construction is complete when, and only when, the final component

[5]The collaborative annotation of a Semantic Web page with semantic interpretations generated by software agents that are beyond the control of its author raises many issues that are outside the scope of this chapter. The annotations might be objectionable to the author or counterproductive to his or her goals; they could be willfully misleading or outright vandalism. While these issues may also arise with the present-day public tagging or bookmarking of sites by users (Breslin et al. 2009), their scale is greatly magnified when the annotations become a central part of the Semantic Web retrieval mechanism rather than merely some user's advisory opinion.

is put in place (even if partial usability is achieved at an earlier stage). By contrast, none or almost none of the features and abilities of an *imperfect* Semantic Web will be fully realized, and it will only imperfectly reflect BLHL's vision; the underlying metaphor is one of growth or evolution, in which even an immature organism is, in an important sense, complete even if not fully functional.

The practical difference between these two views of the development of the Semantic Web is that they lead to different research strategies. And, crucially, we should recognize that the second view is not a lowering but a *raising* of expectations. Why? It reflects the change of view that occurred in computational linguistics and natural language processing in the mid-to-late 1990s, and these fields have been enormously successful since they gave up the too-far-out (or maybe impossible) goal of high-quality knowledge-based semantic interpretation (Hirst 2013) (see Sect. 2). Contemporary NLP and CL have little reliance on symbolic representations of knowledge and of text meaning and far less emphasis on precise results and perfect disambiguation. We have realized that imperfect methods based on statistics and machine learning frequently have great utility; not every linguistic task requires humanlike understanding with 100 % accurate answers; many tasks are highly tolerant of a degree of fuzz and error.

Many other areas of artificial intelligence and knowledge representation came to a similar realization in the last decade or so—just about the time that BLHL's paper was published, but not in time to influence it. In simple terms, BLHL's vision of the Semantic Web is Old School. There needs to be space in the Multilingual Semantic Web for the kinds of imperfect methods now used in NLP and for the textual representations that they imply. In particular, research on vector-based (or tensor-based) distributional semantics (e.g., Turney and Pantel 2010; Clarke 2012; Erk 2012) has reached the point where compositional representations of sentences are now in view (Baroni et al. 2014), and research on distributional methods of semantic relatedness (e.g., Mohammad and Hirst 2006; Hirst and Mohammad 2011) is being extended to cross-lingual methods (e.g., Mohammad et al. 2007; Kennedy and Hirst 2012).

Distributional representations don't meet the "well-defined meaning" criterion of being overtly precise and unambiguous. But it's exactly because of this that they also offer hints of the possibility of reader-centered views of the Semantic Web. Broad distributional representations of a user's search goal, possibly further refined by specific knowledge of other aspects of the user, may match representations of Semantic Web pages that would not be matched by a more precise, symbolic representation of the same goal.

Nonetheless, this can work only if there is agreement on how these representations are constructed from text, including the corpora from which the distributional data are derived. We can envision the development of some kind of standardized lexical or ontolexical vector representation and principles of composition and, moreover, a method of extending the representation across languages. In particular, taking the matter of near-synonymy across languages seriously, we would require that cross-lingual near-synonyms have recognizably similar representations, and hence cross-lingual sentence paraphrases would too.

We should expect to see symbolic representations of textual data increasingly pushed to one side as monolingual and cross-lingual methods are further developed in distributional semantics and semantic relatedness (and a few Semantic Web researchers have already begun some very preliminary investigations Nováček et al. 2011; Freitas et al. 2013). I say this with some caution, as the kind of compositional distributional semantics that could represent phrase and sentence meaning, not just word meaning, and could support useful inference is still at a very early stage of development (e.g., Mitchell and Lapata 2010; Erk 2013; Baroni et al. 2014). In particular, there is no hint yet of a theory of inference for these representations. The whole enterprise might yet fail. But even if this turns out to be so, the broader point remains—that the future of semantic representations for the Multilingual Semantic Web is likely to lie in imperfect nonsymbolic methods that work well enough in practice for most situations.

Acknowledgements This work was supported financially by the Natural Sciences and Engineering Research Council of Canada. For helpful comments, I am grateful to Lars Borin, Philipp Cimiano, Nadia Talent, the anonymous reviewers, and the participants of the Dagstuhl Seminar on the Multilingual Semantic Web.

References

Adamska-Sałaciak, A. (2013). Equivalence, synonymy, and sameness of meaning in a bilingual dictionary. *International Journal of Lexicography, 26*(3), 329–345. doi:10.1093/ijl/ect016.

Baroni, M., Bernardi, R., & Zamparelli, R. (2014). Frege in space: A program for compositional distributional semantics. *Linguistic Issues in Language Technology, 9*(6) (February 2014).

Berners-Lee, T. (2009). The next Web. In *TED Conference*, Long Beach, CA. www.ted.com/talks/tim_berners_lee_on_the_next_web.html.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001, May). The Semantic Web. *Scientific American, 284*(5), 34–43.

Borin, L. (2012). Core vocabulary: A useful but mystical concept in some kinds of linguistics. In D. Santos, K. Lindén, & W. Ng'ang'a (Eds.), *Shall we play the Festschrift game?* (pp. 53–65). Berlin: Springer. doi:10.1007/978-3-642-30773-7_6.

Breslin, J. G., Passant, A., & Decker, S. (2009). *The social Semantic Web*. Berlin: Springer. doi:10.1007/978-3-642-01172-6.

Cimiano, P., Buitelaar, P., McCrae, J., & Sintek, M. (2011). LexInfo: A declarative model for the lexicon–ontology interface. *Journal of Web Semantics, 9*(1), 29–51. doi:10.1016/j.websem.2010.11.001.

Cimiano, P., Unger, C., & McCrae, J. (2014). *Ontology-based interpretation of natural language*. San Rafael: Morgan & Claypool Publishers.

Clarke, D. (2012). A context-theoretic framework for compositionality in distributional semantics. *Computational Linguistics, 38*(1), 41–71. doi:10.1162/COLI_a_00084.

Edmonds, P., & Hirst, G. (2002). Near-synonymy and lexical choice. *Computational Linguistics, 28*(2), 105–144. doi:10.1162/089120102760173625.

Erk, K. (2012). Vector space models of word meaning and phrase meaning: A survey. *Language and Linguistics Compass, 6*(10), 635–653. doi:10.1002/lnco.362.

Erk, K. (2013). Towards a semantics for distributional representations. In *Proceedings, 10th International Conference on Computational Semantics (IWCS-2013)*, Potsdam. www.aclweb.org/anthology/W13-0109.

Farrell, R. B. (1977). *German synonyms*. Cambridge: Cambridge University Press.

Fish, S. (1980). *Is there a text in this class? The authority of interpretive communities*. Cambridge: Harvard University Press.

Freitas, A., O'Riain, S., & Curry, E. (2013). A distributional semantic search infrastructure for linked dataspaces. In *The Semantic Web: ESWC 2013 Satellite Events. Lecture Notes in Computer Science* (Vol. 7955, pp. 214–218). Berlin: Springer. doi:10.1007/978-3-642-41242-4_27.

Friedman, J., Moran, D. B., & Warren, D. S. (1978). Explicit finite intensional models for PTQ. *American Journal of Computational Linguistics, microfiche 74*, 3–22. www.aclweb.org/anthology/J79-1074

Fujiwara, Y., Isogai, H., & Muroyama, T. (1985). *Hyogen Ruigo Jiten*. Tokyo: Tokyodo Publishing.

Gracia, J., Montiel-Ponsoda, E., Cimiano, P., Gómez-Pérez, A., Buitelaar, P., & McCrae, J. (2012). Challenges for the multilingual Web of data. *Journal of Web Semantics, 11*, 63–71. doi:10.1016/j.websem.2011.09.001

Hirst, G. (2007). Views of text-meaning in computational linguistics: Past, present, and future. In G. Dodig Crnkovic & S. Stuart (Eds.), *Computation, information, cognition — The Nexus and the Liminal* (pp. 270–279). Newcastle: Cambridge Scholars Publishing. ftp.cs.toronto.edu/pub/gh/Hirst-ECAPbook-2007.pdf.

Hirst, G. (2008). The future of text-meaning in computational linguistics. In P. Sojka, A. Horák, I. Kopeček, & K. Pala (Eds.), *Proceedings, 11th International Conference on Text, Speech and Dialogue (TSD 2008). Lecture Notes in Artificial Intelligence* (Vol. 5246, pp. 1–9). Berlin: Springer. doi:10.1007/978-3-540-87391-4_1.

Hirst, G. (2009a). Ontology and the lexicon. In S. Staab & R. Studer (Eds.), *Handbook on ontologies. International Handbooks on Information Systems* (2nd ed., pp. 269–292). Berlin: Springer. doi:/10.1007/978-3-540-92673-3_12.

Hirst, G. (2009b, July). Limitations of the philosophy of language understanding implicit in computational linguistics. *Proceedings, 7th European Conference on Computing and Philosophy*, Barcelona (pp. 108–109). ftp.cs.toronto.edu/pub/gh/Hirst-ECAP-2009.pdf.

Hirst, G. (2013). Computational linguistics. In K. Allan (Ed.), *The Oxford handbook of the history of linguistics*. Oxford: Oxford University Press.

Hirst, G., & Mohammad, S. (2011). Semantic distance measures with distributional profiles of coarse-grained concepts. In A. Mehler, K. U. Kühnberger, H. Lobin, H. Lüngen, A. Storrer, & A. Witt (Eds.), *Modeling, learning, and processing of text technological data structures. Studies in Computational Intelligence Series* (Vol. 370, pp. 61–79). Berlin: Springer. doi:10.1007/978-3-642-22613-7_4.

Hirst, G., & Ryan, M. (1992). Mixed-depth representations for natural language text. In P. S. Jacobs (Ed.), *Text-based intelligent systems* (pp. 59–82). Hillsdale, NJ: Lawrence Erlbaum Associates. ftp.cs.toronto.edu/pub/gh/Hirst+Ryan-92.pdf.

Hirst, G., Talent, N., & Scharf, S. (2013, 27 May). Detecting semantic overlap and discovering precedents in the biodiversity research literature. In *Proceedings of the First International Workshop on Semantics for Biodiversity (S4BioDiv)* (CEUR Workshop Proceedings, Vol. 979), *10th Extended Semantic Web Conference* (ESWC-2013), Montpellier, France. ceur-ws.org/Vol-979/.

Hjelmslev, L. (1961). *Prolegomena to a theory of language* (rev. ed.). (F. J. Whitfield, Trans.). Madison: University of Wisconsin Press. (Originally published as *Omkring sprogteoriens grundlæggelse*, 1943.)

Inkpen, D., & Hirst, G. (2006). Building and using a lexical knowledge-base of near-synonym differences. *Computational Linguistics, 32*(2), 223–262. www.aclweb.org/anthology/J06-2003

Kennedy, A., & Hirst, G. (2012, December). Measuring semantic relatedness across languages. In *Proceedings, xLiTe: Cross-Lingual Technologies Workshop at the Neural Information Processing Systems Conference*, Lake Tahoe, NV. ftp.cs.toronto.edu/pub/gh/Hirst-ECAP-2009.pdf.

McCrae, J., Aguado-de-Cea, G., Buitelaar, P., Cimiano, P. Declerck, T. Gómez-Pérez, A., et al. (2012). Interchanging lexical resources on the Semantic Web. *Language Resources and Evaluation, 46*(4), 701–719. doi:10.1007/s10579-012-9182-3.

Mitchell, J., & Lapata, M. (2010). Composition in distributional models of semantics. *Cognitive Science, 34*(8), 1388–1429. doi:10.1111/j.1551-6709.2010.01106.x.

Mohammad, S., Gurevych, I., Hirst, G., & Zesch, T. (2007). Cross-lingual distributional profiles of concepts for measuring semantic distance. In *2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL 2007)*, Prague (pp. 571–580). www.aclweb.org/anthology/D07-1060.

Mohammad, S., & Hirst, G. (2006, July). Distributional measures of concept-distance: A task-oriented evaluation. In *Proceedings, 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*, Sydney, Australia (pp. 35–43). www.aclweb.org/anthology/W06-1605.

Montague, R. (1974). *Formal philosophy.* New Haven: Yale University Press.

Nováček, V., Handschuh, S., & Decker, S. (2011). Getting the meaning right: A complementary distributional layer for the web semantics. In *Proceedings, 10th International Semantic Web Conference (ISWC-2011)* (Vol. 1, pp. 504–519). *Lecture Notes in Computer Science*, Vol. 7031. Berlin: Springer. doi:10.1007/978-3-642-25073-6_32.

Reddy, M. J. (1979). The conduit metaphor: A case of frame conflict in our language about language. In A. Ortony (Ed.), *Metaphor and thought* (pp. 284–324). Oxford: Oxford University Press. [Reprinted unchanged in the second edition, 1993, pp. 164–201.]

Schogt, H. G. (1976). *Sémantique synchronique: synonymie, homonymie, polysémie.* Toronto: University of Toronto Press.

Schogt, H. G. (1988). *Linguistics, literary analysis, and literary translation.* Toronto: University of Toronto Press.

Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research, 37*, 141–188. doi:10.1613/jair.2934.

Wilks, Y. (2009). Ontotherapy, or how to stop worrying about what there is. In N. Nicolov, G. Angelova, & R. Mitkov (Eds.), *Recent advances in natural language processing V* (pp. 1–20). Amsterdam: John Benjamins.