Near-synonymy and the structure of lexical knowledge

Graeme Hirst

Department of Computer Science
University of Toronto
Toronto, Ontario
Canada M5S 1A4
gh@cs.toronto.edu

1 Introduction

Plesionyms, or **near-synonyms**, are words that are almost synonyms, but not quite. The need to deal adequately with plesionymy in tasks such as lexical choice is the basis for two alternatives to conventional models of lexical knowledge: a Saussurean approach and a prototype-theory approach. In this paper, I will discuss these approaches, showing that the latter is troublesome but the former is likely to succeed.

2 A simplistic model of lexical knowledge

Here's the starting point. It looks like a straw man, but it has endured well, having been constructed around a stout cement post.[1]

Any system for understanding or generating natural language needs both a lexicon of the words in the language or languages in question and some knowledge of the world, including a taxonomic hierarchy of concepts. A common view of the relationship between these two structures is that elements of the lexicon hang off the taxonomic hierarchy. That is, the node for any concept in the hierarchy that has a name in the given language—some will have, some won't—will incorporate the lexical entry for that word (or, more precisely, for that word sense). If two or more words denote the same concept, all will be included; if a word is ambiguous, its different senses appear at their respective nodes. And conversely, it is assumed that, for each word sense in the lexicon, there is a node (or ensemble of nodes) somewhere in the hierarchy for it to be attached to; otherwise, the system simply doesn't 'understand' that word sense. Thus, the taxonomic hierarchy, augmented by a lexicon, could look like the simplified fragment shown in Figure 1.

In this view, then, the task of understanding a word in a sentence is to find (presumably by means of an index into the hierarchy) the node or nodes to which it is attached, disambiguate if necessary, and add the result to the structure that is being built to represent the sentence. Conversely, the task of choosing words in natural language generation from a conceptual structure is to find a suitable set of words that 'cover' the structure and assemble them into a sentence in accordance with the syntactic and pragmatic rules of the language (Miezitis 1988; Nogier and Zock 1992; Stede 1993, to appear).

[1]The exposition below is intended as a synthesis, syncretism, or parody of many models found in the literature, not necessarily faithful to any particular one. For examples, see the papers in Evens 1988 (especially Sowa's) and in Pustejovsky and Bergler 1992 (especially those by Nirenburg and Levin, Sowa, and Burkert and Forster). The cement post itself is perhaps Kay 1971.

3 Plesionyms

True synonymy, as simplistically illustrated in Figure 1, is quite rare. It is limited mostly to technical terms (*distichous, two-ranked; groundhog, woodchuck*) and groups of words that differ only in collocational properties, or the like. More frequently, words that are close in meaning are plesionyms—not fully inter-substitutable but varying in their shades of denotation, connotation, implicature, emphasis, or register (DiMarco, Hirst, and Stede 1993, adapting the definitions of Cruse 1986). For example, *lie, falsehood, untruth, fib,* and *misrepresentation* all mean a statement that does not conform to the truth. But a lie is a deliberate attempt to deceive that is a flat contradiction of the truth, whereas a misrepresentation may be more indirect, as by misplacement of emphasis, an untruth might be told merely out of ignorance, and a fib is deliberate but relatively trivial, possibly told to save one's own or another's face (Gove 1984). Moreover, *fib* is an informal, childish term, while *falsehood* is quite formal, and *untruth* can be used euphemistically to avoid some of the derogatory implications of some of the other terms (Gove 1984; compare Coleman and Kay 1981). Table 1 shows a few of the ways in which plesionyms may differ. Often, plesionyms will differ in several ways at once. Some of the examples in the table will be explained below.

It can be difficult even for native speakers of a language to command the differences between plesionyms well enough to use them with invariable precision, or to articulate those differences even when they are known. Consequently, many reference books are published to help in that task (*e.g.,* Bailly 1970; Bénac 1956; Gove 1984; Hayakawa 1968; Room 1985; Urdang 1992). DiMarco and Hirst (1993; DiMarco, Hirst, and Stede 1993) studied such books in order to determine just what kinds of differentiae are adduced between plesionyms.

Table 1: Some of the ways in which plesionyms may differ.

DIFFERENCE	EXAMPLE
Denotation, coarse-grained	*yawl, ketch*
Denotation, fine-grained	*lie, fib, …*
Denotation, fuzzy	*forest, woods*
Emphasis	*foe, enemy*
Implicature	*mislay, lose*
Formality	*drunk, pissed, …*
Attitude of speaker	*skinny, slim*

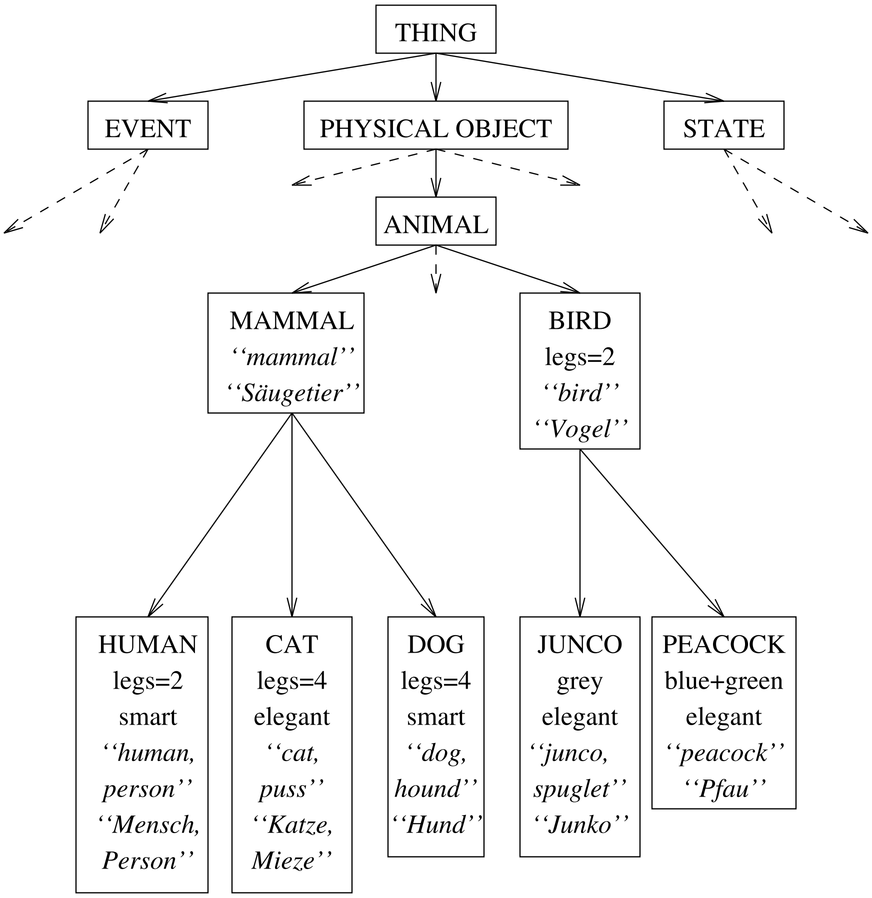

Figure 1: A taxonomic hierarchy, with simplistic lexicons for English and German.

They found that while some differentiae are easily expressed in terms of clear-cut symbolic features (such as CONTINUOUS / INTERMITTENT: *Wine* {*seeped* | *dripped*} *from the barrel*), many are not. Rather, the distinction is a matter of emphasis upon different components of the meaning, or is a vague, ill-defined boundary that might cut across several dimensions simultaneously. For example, the difference between *enemy* and *foe* is that the former stresses antagonism or hatred while the latter stresses active fighting rather than emotional reaction (Gove 1984). The difference between *mislay* and *lose* is that the former implicates an expectation that the missing item will be found and the latter that it won't be; and *mislay* suggests absent-mindedness as a cause, whereas *lose* need not (Hayakawa 1968). The choice between *forest* and *woods* (or *wood*) depends on a complex mixture of size, wildness, and distance from an urban area (Room 1985).[2] Thus the choice between two or more plesionyms cannot necessarily be made merely by a discrimination net–style sequence of ever-finer decisions as to denotation and connotation.

Similar problems occur with lexical transfer in translation: the word in the target language that is closest to that in the source text might be a plesionym rather than an exact synonym. For example, the German word *Wald* is close in meaning to the English word *forest*, but *Wald* can denote a rather smaller and more urban area of trees than *forest*; that is, *Wald* takes in some of the English word *woods* as well, and in some situations, *woods* will be a better translation of *Wald* than *forest*. We can think of *Wald, forest,* and *woods* as a cross-linguistic plesionym group.

4 A Saussurean approach

4.1 Differences are objects

Unfortunately, fine-grained and fuzzy differentiation does not lend itself well to the taxonomic model of the lexicon that we described in Section 2 above. In such a model, each member of a group of plesionyms must be represented as a separate concept (or group of concepts), and, except perhaps

[2]Notice all the hedges in this explanation of the difference: "A 'wood' is smaller than a 'forest', is not so primitive, and is usually nearer to civilization. This means that a 'forest' is fairly extensive, is to some extent wild, and on the whole not near large towns or cities. In addition, a 'forest' often has game or wild animals in it, which a 'wood' does not, apart from the standard quota of regular rural denizens such as rabbits, foxes and birds of various kinds …" (Room 1985, p. 270).

in the case of straightforward symbolic features such as degree of formality, this is not easy or natural. Even simple cases can lead to a multiplicity of concepts that are awkwardly language-dependent. For example, for the *fib* group of plesionyms, we would have to be able to define separate concepts in the hierarchy for UNTRUE-ASSERTION, divided into ACCIDENTAL-UNTRUTH for *untruth*, and DELIBERATE-UNTRUTH, in turn divided into DIRECT-DELIBERATE-UNTRUTH for *lie*, INDIRECT-DELIBERATE-UNTRUTH for *misrepresentation*, and SMALL-FACE-SAVING-DELIBERATE-UNTRUTH for *fib*. And so on, not only for English, but also for the exact divisions made by every other language that is known to the system.[3]

But the transition from the concepts in the knowledge base to the words that denote them has to be made somewhere. The proposal here is that it should be higher rather than lower. That is, the conceptual hierarchy should be fairly coarse-grained—in effect, it should record only relatively language-independent concepts—and the fine tuning, including differentiation between plesionyms, can then be done in the lexical entries for each separate language.[4]

The reader will possibly have recognized that the preceding discussion is implicitly Saussurean; so let us make it explicit. For Saussure, "although in general a difference presupposes positive terms between which the difference holds, in a language there are only differences, *and no positive terms*" (1916, p. 166). Taken to this extreme, the position is paradoxical; but we need only milder versions of the same idea—for example, Clark's (1992) Principle of Contrast: "Every two [linguistic] forms contrast in meaning" (p. 172).

The idea, then, is that we recognize differences or contrasts as objects in their own right that can be represented and reasoned about (and, for that matter, taxonomized and differentiated). We can then use these objects in describing the nuances of meaning that distinguish plesionyms from one another and in choosing among them. In other words, our lexicon will no longer be defined simply in positive terms; rather, each concept will now contain information on differentiation of any set of plesionyms that is mapped to it. This information may, nonetheless, be expressed in the vocabulary of the hierarchy itself, with connectives and operators based upon the classes of lexical differentiation described by DiMarco, Hirst, and Stede (1993).

4.2 Differences between concepts

In order to see how we can treat differences as first-class objects, we'll first consider the simpler case of differences between concepts before we look at sub-conceptual differences between plesionyms. In particular, we'll consider concepts that are close to one another in the taxonomic hierarchy and are **structurally alignable** (Gentner and Markman 1994)—that is, the two concepts have many attributes in common, albeit with possibly different values for those attributes. (Thus the concepts of *bus* and *train* are structurally alignable; those of *canary* and *shopping mall* are not.)

For example, the difference between a *yawl* and a *ketch* is that, while both are two-masted sailing yachts, a yawl has the mizzenmast set aft of the rudderpost instead of forward.[5] If our knowledge of various types of sailing vessels is kept in a taxonomic hierarchy, perhaps as shown in Figure 2, then this difference is a set of contrasting attribute-value pairs, where each attribute either takes on different values for each object or is missing entirely from one of the objects. In effect, it is a proposition that we might write like this:

(1)
```
Diff (yawl / ketch) =
      [location (mizzen) =
       aft-of-rudderpost /
       forward-of-rudderpost]
```
[6]

'The difference between a yawl and a ketch is that the location of the mizzenmast is aft-of-rudderpost in the former and forward-of-rudderpost in the latter'

I allow `Diff` to take as arguments both the symbols for concepts in the taxonomic hierarchy and the words associated with them. So for the previous example, we could equivalently have written `Diff ("Besankutter" / "Ketsch")` or (cross-linguistically) `Diff ("Besankutter" / "ketch")` or even `Diff ("Besankutter" / ketch)`.

Clearly, these differences can be computed from the taxonomic hierarchy by collecting the attributes specified in each node along the paths from the two objects to their most specific subsumer (which, in this example, is their immediate parent). If there are any attributes along the paths that are identical, they are deleted from the set. Thus, in the hierarchy of animals shown in Figure 1, it just happens that humans and birds, although taxonomically quite distinct, both have two legs, and hence we have:

(2)
```
Diff (human / junco) =
      [smart / (grey, elegant)]
```
[7]

'The differences between a human and a junco are that only the former is smart and only the latter is grey and elegant'

I conjecture that this kind of computation and reification of differences has some claim to psychological reality. Gentner and Markman (1994) have shown that people can very easily articulate the differences between two atructurally alignable

[3]Some systems have indeed taken this approach, *e.g.,* Emele, Heid, Momma, and Zajac 1992.

[4]I thus sidestep the long-debated question of exactly what degree of inter-substitutability should count as synonymy or near-synonymy (Egan 1942; Sparck Jones 1986; Cruse 1986). For this work, synonymy and near-synonymy arise at the point in the conceptual network at which a (language-independent) concept diverges into the set of (language-dependent) lexical entries for the words that, in one way or another, denote that concept. That is, the groups of words among which we need to discriminate are exactly the groups of words (in each language of interest) that correspond to each single concept in the taxonomic hierarchy of the knowledge base of the system. For convenience, I shall refer to each group as a set of plesionyms or near-synonyms, but I intend by these labels no theoretical import beyond that of this operational definition.

[5]This example was suggested to me by Laurence Urdang.

[6]My notation should be reasonably transparent. `Diff` takes two arguments, separated by a slash. Its value is a list of differences (only one in this example). Where the two arguments have different values for a shared attribute, I write the difference as the attribute name followed by an equals-sign and the two values separated by a slash, as seen here. Where one argument has an attribute that the other entirely lacks, I will write the attribute–value pair on one side of the slash and ∅ on the other.

[7]I write this as a convenient abbreviation for `[smart / ∅; ∅ / grey; ∅ / elegant]`.

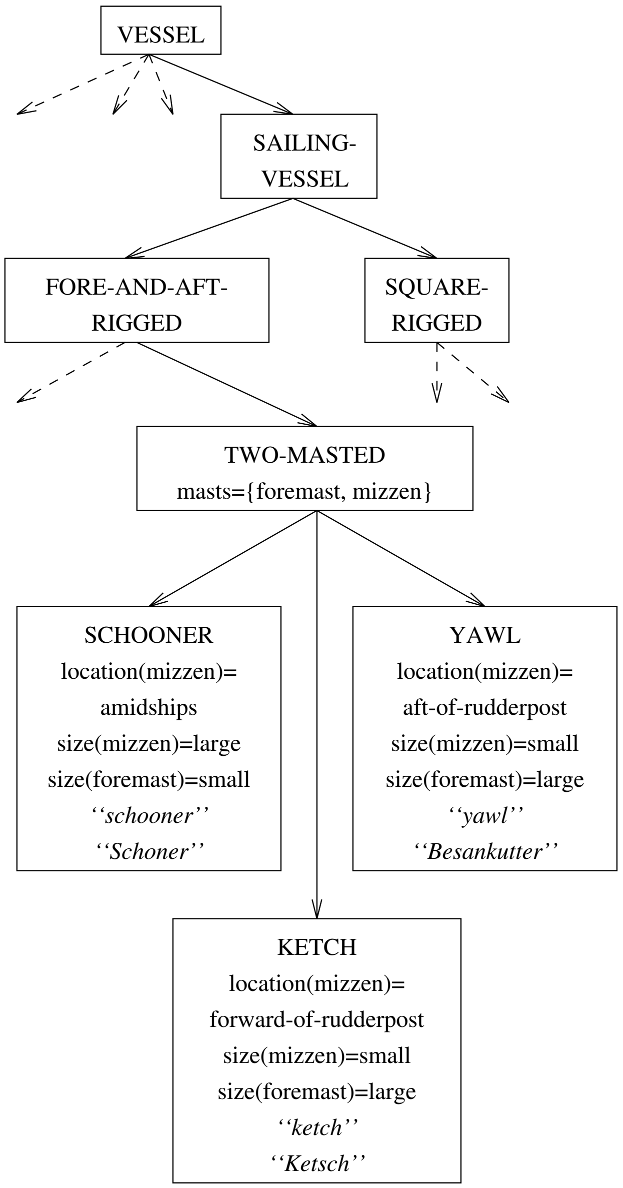

Figure 2: A fragment of a field guide to yachts. The conceptual and English data are from Urdang 1992; the German data are from *Langenscheidt's New College German Dictionary* (New York: Langenscheidt, 1991).

concepts. If people have the ability to compute differences, we would expect that they sometimes keep the results—that is, that they have explicit, pre-compiled knowledge of differences between closely related concepts and closely related words. Research such as that of Gentner and Markman is consistent with this conjecture, but I am not aware of any research that explicitly tests it.

4.3 Comparing differences

I alluded earlier to the problem of lexical transfer in machine translation when the nearest words in the target language are but plesionyms of the source word, each a little different in meaning from it. To decide which target word might be best in a particular situation, we need the notion of finding the **smallest difference**—that is, in general, we need to be able to compare differences between pairs of words (whether in the same language or different languages). The implication is that in MT we will often need to explicitly represent differences between cross-linguistic plesionyms. Indeed, we should expect this, for exactly the same reason that we need differentiation and usage notes in a single language.

A comparison of conceptual differences will need to incorporate several factors. For example, if D_1 and D_2 are differences, we might write $D_1 < D_2$ if there is an ordering on the attribute-values that appear in them:

(3) `[legs = 2 / 4] < [legs = 2 / 6]`
'The difference between a two-legged object and a four-legged object is smaller than the difference between a two-legged object and a six-legged object.'

We might also say that $D_1 < D_2$ if D_2 extends D_1 with additional differences:

(4) `[legs = 2 / 4] <`
`[legs = 2 / 4; smart / ∅]`
'The difference between a two-legged object and a four-legged object is smaller than the difference between a smart two-legged object and a four-legged object.'

Clearly, two differences need not be comparable at all. There is no relationship apparent between, say, `[legs = 2 / 4]` and `[mainsail = quadrilateral / triangular]`.

4.4 Sub-conceptual differences

Now, if our differentiae used only the conceptual information of the taxonomic hierarchy, we would gain little. We need to also be able to talk about fine-grained sub-conceptual differences; for example, `Diff ("meddlesome" / "prying")`, where both words map to the same concept in the hierarchy—`interfering`, let's say—and hence, by our definition above, the difference is empty. The proposal is that such plesionymic differences must be explicitly incorporated into the lexical entries of the words. That is, we have in effect, a 'usage note' that describes how the words differ and when one or the other might be more appropriate. Such 'usage notes' need not be propositional—in fact, they would be a grab-bag of tendencies, emphases, implicatures, microfeatures (both connotational and denotational), discrimination nets (where possible) or other decision procedures, exemplars, and so on. We thus acknowledge the Saussurean notion that meaning at this level is expressible only in terms of differences.

So that the differences adduced by such usage notes can be first-class objects that can be reasoned about and compared with one another, just like our conceptual-level differences, I extend the definition of `Diff` so that it returns the appropriate usage note when its arguments are isoconceptual words. Moreover, the procedures that operate upon differences must be able to work on all types of differences, including mixtures of different types.

Developing suitable representations and procedures for these usage notes is the goal of our present research.

5 The prototype-theory approach

One likely kind of usage note that does not fit the Saussurean notion well is one that is based on **exemplars**. For example, `Diff ("forest" / "woods")` might best be expressed not in comparative terms but rather by presenting a canonical example of each.

The idea that some plesionym groups can be differentiated by reference to exemplars is reminiscent of **prototype theory**. Prototype theorists, best exemplified by Lakoff (1987), reject the notion that concepts can be organized by necessary and sufficient conditions into a taxonomic hierarchy such as that of Figure 1. Instead, knowledge is organized into categories "by means of" (p. 68) **idealized cognitive models** (ICMs)—gestalts that include a schema-like propositional structure (à la Fillmore 1982, perhaps as developed by Barsalou 1992), an image-schematic structure, and metaphoric and metonymic mappings. (Compare the lexical entries of Allan (1990, 1992).) Membership in a category may be based on **similarity to a prototype**, and may be a matter of degree.

In this framework, Lakoff (pp. 131*ff*) analyzes word pairs such as *thrifty* / *stingy* as, in effect, competing idealized cognitive models.[8] Thus, in this analysis, *Ross is thrifty* and *Ross is stingy* can denote the same state of affairs, or have the same truth conditions; but they differ in that the former evokes (or 'pragmatically presupposes') an idealized cognitive model in which spending as little money as possible is thought to be good, whereas the latter evokes an ICM in which it is thought to be bad. And with a sentence such as *Ross isn't stingy, he's thrifty*, the speaker is explicitly rejecting one ICM in favor of the another.

Now, this analysis is simplistic—*thrifty* and *stingy* differ in more than just the speaker's expressed attitude[9]—but a coherent position can be derived from it: that (except in the rare case of completely inter-substitutable synonyms) **any** contrast between two or more words is a contrast between ICMs. This would be an explicit rejection of the Saussurean approach, as ICMs, being gestalts, are (I assume) defined solely in positive terms. And it would be a claim that, no matter how close two plesionyms are, they nonetheless have distinct ICMs.[10]

As a psychological claim, such a position is quite probably wrong. As we observed in Section 4.2 above, it is likely that people have explicit mental representations of differences between at least some near-synonyms as part of their knowledge of the meanings of the words. But this seems to be at odds with the orthodox prototype-theory view of the lexicon.

Nevertheless, a tenable view of some aspects of plesionymy can indeed be developed within the framework of prototype theory. I suggest that plesionyms often do represent the same concept or ICM, but they exhibit differing prototypes of the concept or denote variations in degrees of exemplification or membership of that concept.

Consider again, for example, the case of *forest* and *woods*. While each has, presumably, a different prototype, it need not follow that they have different ICMs. For if we were to start to describe the ICM for each, we would find that they were virtually identical, differing only by the position of the prototype in the 'space' of size, wildness, and so on (see footnote 2). We could equivalently regard the two words as sharing a single ICM within which each word has a distinct prototype. A similar argument can be made for other groups of plesionyms in which one word shades into another and their differences lie in matters of degree: *strait, sound; stingy, miserly; mist, fog;* and so on. In other words, we represent such plesionym groups by a single ICM, or concept, or however we wish to characterize it, and represent each word in the group by a separate exemplar or prototype. When we need to choose one of the words in the group to refer to some particular object, we pick the one whose prototype is most similar to the object.[11]

Unfortunately, like prototype theory in general, this is easy to describe but hard to make computationally precise and useful. Certainly, no such model of the lexicon has yet been developed within prototype theory (though the work of Allan (1992) is a small first step). (Indeed, on some interpretations of prototype theory, such a model would be thought to be impossible in principle, or possible only if we are "fortunate" (Lakoff 1987, p. 287).) To the extent that prototypes are propositional in nature, and to the extent that we can develop a computational theory of similarity for comparing other objects to such propositional prototypes, we could perhaps incorporate them into a Saussurean model of the lexicon as described in section 4 above, but in so doing we would be throwing away most of what prototype theory stands for.[12]

6 Conclusion

The fine nuances of many kinds that plesionyms exhibit make them difficult, if not impossible, to represent adequately in conventional, symbolic, taxonomic models of the lexicon in which words are denoted in strictly positive terms. Some plesionyms seem to be best defined by means of different exemplars or prototypes of the same concept. Prototype theory, however, offers little or no assistance in building models of the lexicon that can be applied computationally. Rather, a multifaceted representation seems to be required in which differences and contrasts are themselves objects of various types that can be adduced as part of the definition of a group of plesionyms. Our goal is the development of such a representation.

[8] Lakoff's analysis is based largely on that of Fillmore (1982, 1985).

[9] Specifically, *thrifty* implies a careful use of limited resources to maximal utility, whereas *stingy* implies selfishness, greed, and a close-fistedness beyond any that circumstances necessitate. Gove (1984) doesn't regard them even as members of the same plesionym group.

[10] It's not clear whether Lakoff actually intended this to be inferred from his discussion. His prolix, holographic style of writing makes it hard for one to be certain of his precise position here.

[11] Obviously, I don't claim that all plesionym groups can be analyzed this way. In particular, the effect of **euphemism** (and **dysphemism**) is usually achieved exactly by means of choosing a word or phrase with a different ICM in order to avoid (or intensify) the unpleasant imagery or associations in the ICM of the original, plesionymous word. This is particularly evident in the often-parodied penchant of people of progressive politics to coin ponderous euphemisms in the hope of changing societal attitudes to personal attributes thought to be inappropriately deprecated.

[12] Grandy (1992) does just this when, in defining the idea of a semantic field, he includes prototypes in his taxonomically organized "contrast sets".

Acknowledgements

I am grateful to Chrysanne DiMarco for many helpful discussions. The preparation of this paper was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

References

Allan, Keith (1990). "Some English terms of insult involving sex organs: Evidence of a pragmatic driver for semantics." In: Tsohatzidis, Savas L. (editor). *Meanings and prototypes: Studies in linguistic categorization.* London: Routledge, 159–194.

Allan, Keith (1992). " 'Something that rhymes with rich'." In Lehrer and Kittay 1992, 355–374.

Bailly, René (1970). *Dictionnaire des synonymes de la langue française*. Paris: Librarie Larousse.

Barsalou, Lawrence W. (1992). "Frames, concepts, and conceptual fields." In Lehrer and Kittay 1992, 21–74.

Bénac, Henri (1956). *Dictionnaire des synonymes*. Paris: Librarie Hachette.

Burkert, Gerrit and Forster, Peter (1992). "Representation of semantic knowledge with term subsumption languages." In Pustejovsky and Bergler 1992, 75–85.

Clark, Eve V. (1992). "Conventionality and contrast: Pragmatic principles with lexical consequences." In Lehrer and Kittay 1992, 171–188.

Coleman, Linda and Kay, Paul (1981). "Prototype semantics: The English word *lie*." *Language*, **57**(1), March 1981, 26–44.

Cruse, D.A. (1986). *Lexical semantics*. Cambridge University Press.

DiMarco, Chrysanne and Hirst, Graeme (1993). "Usage notes as the basis for a representation of near-synonymy for lexical choice." *Proceedings, 9th annual conference of the University of Waterloo Centre for the New Oxford English Dictionary and Text Research*, Oxford, September 1993, 33–43.

DiMarco, Chrysanne; Hirst, Graeme; and Stede, Manfred (1993). "The semantic and stylistic differentiation of synonyms and near-synonyms." *AAAI Spring Symposium on Building Lexicons for Machine Translation*, Stanford, CA, March 1993, 114–121.

Egan, Rose F. (1942). "Survey of the history of English synonymy" and "*Synonym:* Analysis and definition." Reprinted in Gove 1984, pp. 5a–25a.

Emele, Martin; Heid, Ulrich; Momma, Stefan; and Zajac, Rémi (1992). "Interactions between linguistic constraints: Procedural vs. declarative approaches". *Machine Translation*, **7**(1–2), 1992, 61–98.

Evens, Martha Walton (editor) (1988). *Relational models of the lexicon: Representing knowledge in semantic networks.* Cambridge University Press.

Fillmore, Charles J. (1982). "Frame semantics." In *Linguistics in the morning calm: Selected papers from SICOL-1981*, Linguistic Society of Korea, Seoul: Hanshin Publishing Company, 1982, 111–137.

Fillmore, Charles J. (1985). "Frames and the semantics of understanding." *Quaderni di semantica*, **6**(2), 222–254.

Gentner, Dedre and Markman, Arthur B. (1994). "Structural alignment in comparison: No difference without similarity." *Psychological science*, **5**(3), May 1994, 152–158.

Gove, Philip (editor) (1984). *Webster's new dictionary of synonyms*. Springfield, Mass.: Merriam-Webster.

Grandy, Richard E. (1992). "Semantic fields, prototypes, and the lexicon." In Lehrer and Kittay 1992, 103–122.

Hayakawa, S.I., (editor) (1968). *Funk & Wagnall's modern guide to synonyms and related words*. New York: Funk & Wagnall's division of Reader's Digest Books. Reprinted as *Use the right word: Reader's Digest modern guide to synonyms and related words*, Reader's Digest Association, no date. Reprinted as *Choose the right word: A modern guide to synonyms*, New York: Harper & Row (Perennial Library), 1987. Edition for British English (revised by Paul J. Fletcher) published as *Cassell's modern guide to synonyms*, London: Cassell, 1971; republished as *The Penguin modern guide to synonyms and related words*, Penguin, 1987. Second edition (revised by Eugene Ehrlich) published as *Choose the right word: A contemporary guide to selecting the precise word for every situation*, New York: Harper Perennial, 1994.

Kay, Paul (1971). "Taxonomy and semantic contrast." *Language*, **47**(4), December 1971, 866–887.

Lakoff, George (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: The University of Chicago Press.

Lehrer, Adrienne and Kittay, Eva Feder (1992). *Frames, fields, and contrasts: New essays in semantic and lexical organization.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Miezitis, Mara (1988). *Generating lexical options by matching in a knowledge base.* MSc thesis, Department of Computer Science, University of Toronto [published as technical report CSRI-217].

Nirenburg, Sergei and Levin, Lori (1992). "Syntax-driven and ontology-driven lexical semantics." In Pustejovsky and Bergler 1992, 5–20

Nogier, Jean-François and Zock, Michael (1992). "Lexical choice as pattern matching." In: Nagle, Timothy E.; Nagle, Janice A.; Gerholz, Laurie L.; and Eklund, Peter W. (editors) (1992). *Conceptual structures: Current research and practice*, Chichester: Ellis Horwood, 413–435.

Pustejovsky, James and Bergler, Sabine (editors) (1992). *Lexical semantics and knowledge representation* (Lecture notes in artificial intelligence 627). Berlin: Springer-Verlag.

Room, Adrian (1985). *Dictionary of confusing words and meanings.* London: Routledge & Kegan Paul.

Saussure, Ferdinand de (1916). *Cours de linguistique générale*. English translation by Roy Harris as *Course in general linguistics*, 1983, London: Duckworth. [Page numbers cited are those of the second French edition, indicated by marginal notes in this translation.]

Sowa, John (1988). "Using a lexicon of graphs in a semantic interpreter." In Evens 1988, 113–137.

Sowa, John (1992). "Logical structures in the lexicon." In Pustejovsky and Bergler 1992, 39–60.

Sparck Jones, Karen (1986). *Synonymy and semantic classification*. Edinburgh: Edinburgh University Press (Edinburgh information technology series 1).

Stede, Manfred (1993). "Lexical options in multilingual generation from a knowledge base." In: Grote, Brigitte; Roesner, Dietmar; Stede, Manfred; and Weis, U. *From knowledge to language—Three papers on multilingual text generation*. Technical report 93017, Forschungsinstitut für anwendungsorientierte Wissensverarbeitung, Ulm.

Stede, Manfred (to appear). "Lexicalization in natural language generation: A survey." *Artificial Intelligence Review*, in press.

Urdang, Laurence (1992). *Dictionary of differences*, revised edition. London: Bloomsbury Publishing.