A selective annotated bibliography for computational lexical semantics

Ann Copestake
CSLI, Stanford University
aac@csli.stanford.edu
May 1996

This bibliography is not comprehensive or unbiased. Unfortunately it only includes material written in English. My intention was to give someone who was unfamiliar with recent work on computational lexical semantics some pointers, mostly to work which is relatively easily available, either through conventional publication or electronically. Since this bibliography is based on the collection of references I've used in papers, it is distinctly skewed towards my own and related work, though I've tried to add references where it was obvious that there was a big gap. In these cases, I've often only added a single reference to an author's work, and I've often chosen easily accessible papers in preference to more obscure ones, even when the latter are more comprehensive.

Many of these references are not directly about computational treatments of lexical semantics, but I have included them because they motivated a computational approach or provided a good overview of the issues without requiring detailed knowledge of a particular linguistic theory. With respect to formalization, this bibliography is heavily biased toward my own interests: that is the use of constraint-based frameworks incorporating defaults and lexical rules. A number of references to defaults and to lexical rules in unification-based formalisms have been included even though they do not directly concern lexical semantics.

On the whole I have concentrated on work which treats computational lexical semantics from a linguistic perspective: that is, work which motivates semantic representation because of its interaction with syntax, morphology and so on. I have not included many references to the work on representing meaning from the contrasting AI/Knowledge representation tradition, largely because the literature on this is unbounded. However, several papers which take this perspective are in the collections referenced in the first section (especially Saint-Dizier and Viegas (1995) and Pustejovsky and Bergler (1992)).

The bibliography is divided into the following sections:

1. General
2. Event structure
3. Lexical semantics and pragmatics
4. Verbs, syntax-semantics interface
5. Derivational morphology and compounds
6. Polysemy
7. Metaphors, idioms and collocations
8. Formal representation: inheritance and defaults
9. Formal representation: lexical rules
10. Disambiguation
11. Acquisition

Introductory notes give a brief idea of the coverage intended. Some papers appear in more than one section.

An updated version of this bibliography will be available via my home page: `http://www-csli.stanford.edu/users/aac`. Comments and suggestions welcome.

1 General

Levin (1985 (1995)) and Briscoe (1996) are both reviews of the literature which give more detail than is attempted here. Neither is a complete overview: Levin's paper mostly concerns linguistic work on verbal lexical semantics, Briscoe's is largely a summary of some of the work on the ACQUILEX project in the context of other work on lexical semantics. The introductions to Levin and Pinker (1992) and to Saint-Dizier and Viegas (1995) are also useful reviews.

Unfortunately several of the collections listed here were not published until long after the papers which they contain were originally written, and they are thus somewhat dated. Specifically, the papers in Saint-Dizier and Viegas (1995) were written in 1991/1992, Pustejovsky (1993) grew out of a workshop held in 1988, Walker *et al.*(1995) contains papers written in 1985. Pustejovsky and Bergler (1992) was published much more quickly and is essentially the same as the workshop proceedings (SIGLEX Workshop, 1991). Levin and Pinker (1992) contains papers published in *Cognition* (1991, Vol 41, 1–3). All of these collections contain a mixture of computational and theoretical linguistic papers.

Levin (1993) is a descriptive account of verb classes and diathesis alternations which is unique in providing data for over 3000 verbs. Since it also contains an extensive bibliography it is an essential reference. Cruse (1985) is a linguistic textbook which gives a descriptive account of lexical semantics. It is a particularly useful source for terminology and its discussion of polysemy. Pustejovsky and Boguraev (1993) discusses the use of knowledge representation techniques in computational lexical semantics.

Bierwisch and Schreuder (1992), Copestake (1992), Dowty (1979), Jackendoff (1990), Kamp and Rossdeutscher (1994a,b), Mel'čuk and Polguère (1987), Moravcsik (1975) and Pustejovsky (1991,1995) are included here because they describe different approaches to lexical semantics, covering a range of phenomena. Dowty (1979) is particularly important with respect to event structure, but is also a more general account of how lexical semantics could be approached within Montague Grammar.[1] Pustejovsky's work has been extremely influential within recent work on computational lexical semantics: some of his ideas, in particular his notion of *qualia structure* is partially based on Moravcsik (1975). Copestake (1992) is a computational approach which partly consists of a formalization and implementation of some of Pustejovsky's ideas, but also draws on work by several other authors cited here.

Bierwisch, M. and R. Schreuder (1992) 'From concepts to lexical items', *Cognition,* **42**, 23–60.

Briscoe, E.J. (1996) 'Computational lexical semantics in review', Part of FraCaS deliverable, obtainable via `http://www-csli.stanford.edu/users/aac`.

Copestake, A. (1992) 'The representation of lexical semantic information', Doctoral dissertation, University of Sussex, Cognitive Science Research Paper CSRP 280

[1]Obviously I do not accept the characterization of lexical semantics as being in any sort of opposition to formal semantics. I haven't tried to systematically distinguish between informal and formal treatments in this bibliography. Although an overtly formal approach may be stated more precisely and thus be easier to implement, very little of the linguistic work considers anything remotely resembling efficiency (with respect to time or space) or ease of acquisition. So a clearly stated descriptive account may be more useful to a computational linguist if it suggests a formalization using existing tractable techniques.

http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html.
Cruse, D.A. (1986) *Lexical semantics,* Cambridge University Press, Cambridge, UK.
Dowty, D. (1979) *Word meaning in Montague Grammar,* Reidel, Dordrecht.
Jackendoff, R. (1990) *Semantic structures,* MIT Press, Cambridge, MA.
Kamp, H. and A. Rossdeutscher (1994a) 'DRS-construction and lexically driven inference', *Theoretical linguistics.*
Kamp, H. and A. Rossdeutscher (1994b) 'Remarks on lexical structure and DRS construction', *Theoretical linguistics.*
Levin, B. (1985 (1995)) 'Approaches to lexical semantic representation (Lexical semantics in review: an introduction)' in D. Walker, A. Zampolli and N. Calzolari (ed.), *Automating the lexicon: research and practice in a multilingual environment,* Oxford University Press, Oxford, pp. 53–92.
Levin, B. (1993) *Towards a lexical organisation of English verbs,* University of Chicago Press, Chicago.
Levin, B. and S. Pinker (1992) *Lexical and conceptual semantics,* Blackwell, Cambridge MA and Oxford, UK.
Mel'čuk, I. and M. Polguère (1987) 'A formal lexicon in Meaning-Text Theory (or how to do lexica with words)', *Computational Linguistics,* **13:3–4**, 261–275.
Moravcsik, J. (1975) 'Aitia as generative factor in Aristotle's philosophy', *Dialogue,* **14**, 622–636.
Pustejovsky, J. (1991) 'The Generative Lexicon', *Computational Linguistics,* **17.4**, 409–441.
Pustejovsky, J. (ed.) (1993) *Semantics and the lexicon,* Kluwer, Dordrecht.
Pustejovsky, J. (1995) *The Generative Lexicon,* MIT Press, Cambridge, MA.
Pustejovsky, J. and B. Boguraev (1993) 'Lexical knowledge representation and natural language processing', *Artificial Intelligence,* **63**, 193–223.
Pustejovsky, J. and S. Bergler (eds.) (1992) *Lexical semantics and knowledge representation. Proceedings of the first SIGLEX workshop, Berkeley, CA,* Springer-Verlag, Berlin.
Saint-Dizier, P. and E. Viegas (1995) *Computational lexical semantics,* Cambridge University Press, Cambridge, UK.
Walker,D., A. Zampolli and N. Calzolari (1995) *Automating the lexicon: research and practice in a multilingual environment,* Oxford University Press, Oxford.

2 Event structure

There have been various elaborations of event structure beyond the Vendlerian classification into states, processes, achievements and accomplishments, and this section just contains a few pointers to the extensive literature. Dowty (1979) presents an account based on the use of four primitives BE, BECOME, CAUSE and DO. Parson's neo-Davidonian approach uses event variables (unlike Dowty) — this has become a very common technique in computational linguistics. Krifka (1987) provides an account which relates event structure to the structure of nominals, formalizing the parallelism between the bounded/unbounded distinction for events and the cumulative/quantized distinction for nouns. Moens and Steedman (1988) and Pustejovsky (1991) argue for more complex event structures than those assumed by Parsons. Pustejovsky and Busa (1995) extend Pustejovsky (1991) by arguing for a notion of headedness on event structure.

Dowty, D. (1979) *Word meaning in Montague Grammar,* Reidel, Dordrecht, The Netherlands.
Krifka, M. (1987) 'Nominal reference and temporal constitution: towards a semantics of quantity', *Proceedings of the 6th Amsterdam Colloquium,* University of Amsterdam, pp. 153–173.

Moens, M. and M. Steedman (1988) 'Temporal ontology and temporal reference', *Computational Linguistics,* **14**, 15–28.

Parsons, T. (1990) *Events in the semantics of English: a study in subatomic semantics,* MIT Press, Cambridge, Massachusetts.

Pustejovsky, J. (1992) 'The syntax of event structure' in B. Levin and S. Pinker (ed.), *Lexical and conceptual semantics,* Blackwell, Cambridge MA and Oxford, UK.

Pustejovsky, J. and F. Busa (1995) 'Unaccusativity and event composition' in P. Bertinetto, V. Binachi, J. Higginbotham, M. Squartini (ed.), *Temporal reference: aspect and actionality,* Rosenberg and Sellier, Turin.

3 Lexical semantics and pragmatics

I have only attempted to include papers here which give a formal/computational account of pragmatics and explicitly relate it to lexical semantics. The approach of Hobbs *et al.*(1990) explicitly rejects the separation of lexical semantics in favour of a uniform processing mechanism. Lascarides and coauthors, in contrast, assume a more modular approach, which makes use of the approach to discourse representation described by Lascarides and Asher (1993). Kayser and Abir (1995) describe a somewhat different framework, which also makes use of non-monotonic logic.

Hobbs, J.R., M. Stickel, D. Appelt and P. Martin (1990) 'Interpretation as abduction', Technical Note No. 499, Artificial Intelligence Center, SRI International, Menlo Park, CA.

Kayser, D. and H. Abir (1995) 'A non-monotonic approach to lexical semantics' in Saint-Dizier, P. and E. Viegas (ed.), *Computational lexical semantics,* Cambridge University Press, Cambridge, UK, pp. 303–318.

Lascarides, A (1995) 'The pragmatics of word meaning', *Proceedings of the AAAI Spring Symposium on lexical semantics*, Stanford CA.

Lascarides, A. and A. Copestake (1996) 'The pragmatics of word meaning' in M. Simons and T. Galloway (ed.), *Semantics and Linguistic Theory V,* Cornell University, Ithaca, NY, pp. 204–221.
`http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`

Lascarides, A. and N. Asher (1993) 'Temporal interpretation, discourse relations and common sense entailment', *Linguistics and Philosophy,* **16**, 437–493.

Lascarides, A., A. Copestake and E.J. Briscoe (1996, in press) 'Ambiguity and coherence', *Journal of Semantics,* **13:2**.
`http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`

4 Verbs, syntax-semantics interface

Again, these references are merely pointers into an extensive literature. Only Sanfilippo's work is computational.

Davis, A. (1996) 'Linking and the hierarchical lexicon', Doctoral dissertation, Stanford University.

Dowty, D. (1989) 'On the semantic content of the notion 'Thematic Role" in G. Chierchia, B. Partee and R. Turner (ed.), *Property Theory, Type Theory and Natural Language Semantics,* D. Reidel, Dordrecht, The Netherlands, pp. 69–129.

Fillmore, C. (1968) 'The case for case' in E. Bach and T. Harms (ed.), *Universals in linguistic theory,* Holt, Rinehart, Williams, New York, pp. 1–88.

Fillmore, C. and P. Kay (in press) *Construction Grammar,* CSLI Publications.

Goldberg, A. (1995) *Constructions: a construction grammar approach to argument structure,* Chicago University Press.

Levin, B. (1993) *Towards a lexical organisation of English verbs,* University of Chicago Press, Chicago.

Pinker, S (1989) *Learnability and cognition: the acquisition of argument structure,* MIT Press, Cambridge, MA.

Sanfilippo, A. (1992) 'Verbal diathesis, knowledge acquisition, lexicon construction and dictionary compilation', Acquilex-II Working Papers 1: see also Proc. Coling 94.

Sanfilippo, A. (1993) 'LKB encoding of lexical knowledge from machine-readable dictionaries' in E. J. Briscoe, A. Copestake and V. de Paiva (ed.), *Inheritance, defaults and the lexicon,* Cambridge University Press, Cambridge, UK, pp. 190–222.

Talmy, L. (1985) 'Lexicalization patterns' in T. Shopen (ed.), *Language typology and syntactic description,* Cambridge University Press, Cambridge, UK.

Zaenen, A. (1993) 'Unaccusativity in Dutch: the interface of syntax and lexical semantics' in J. Pustejovsky (ed.), *Semantics and the lexicon,* Kluwer, Dordrecht.

5 Derivational morphology and compounds

I haven't cross-classified all the papers which could appear in this section. Derivational morphology and regular polysemy share many properties and pose similar problems for a computational account: in particular semi-productivity, semi-regularity of meaning (i.e. lexicalised senses and derived forms often have meanings which are not fully predictable), problems in formalization of the semantics, and technical issues about formalization with lexical rules in a unification-based framework. So some papers relevant to derivational morphology are in the sections on polysemy and lexical rules. Compounds pose even greater problems with respect to predictability of meaning, and it is often claimed that their meaning must be pragmatically determined — see also section on pragmatics.

Bauer (1983) is a textbook, but it contains several interesting ideas, about productivity and compounds in particular, and a useful summary of various classes of English affixes. Levi's (1978) book is a standard reference on nominal compounds. Wu (1990) and Johnston *et al.*(1995) describe (somewhat different) computational approaches to processing compounds (which still remains an intractable problem — there's a fairly large literature on this which I hope the cited papers provide adequate references to). Light (1996) is actually a description of how to acquire some lexical semantic information using derivational morphology as a cue, but it's also useful for its illustration of the formalization of some aspects of the meaning of derived forms. Riehemann (1994) combines formalization with a discussion of productivity motivated by extensive corpus work. Barker (1996) also discusses productivity.

Barker, C. (1996) 'Episodic *-ee* in English: thematic relations and new word formation' in M. Simons and T. Galloway (ed.), *Semantics and Linguistic Theory V,* Cornell University, Ithaca, NY:, pp. 1–18.

Bauer, L. (1983) *English word-formation,* Cambridge University Press, Cambridge, UK.

Johnston, M., B. Boguraev and J. Pustejovsky (1995) 'The acquisition and interpretation of complex nominals', *Proceedings of the AAAI Spring Symposium on representation and acquisition of lexical knowledge,* Stanford, CA.
`http://www.cs.brandeis.edu/~johnston/home.html`

Levi, J. (1978) *The syntax and semantics of complex nominals,* Academic Press, New York.

Light, M. (1996) 'Morphological cues for lexical semantics', *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics (ACL-96),* Santa Cruz, CA.
`http://www.sfs.nphil.uni-tuebingen.de/~light/morphcues.ps.gz`

Riehemann, S. (1994) 'Morphology and the hierarchical lexicon', Stanford University
`http://hpsg.stanford.edu/hpsg/papers.html`.

Wu, D. (1990) 'Probabilistic unification-based integration of syntactic and semantic preferences for nominal compounds', *Proceedings of the 13th International Conference on Computational Linguistics (Coling90),* Helsinki, pp. 413–418.

6 Polysemy

The coverage in this section is somewhat more extensive, although I have omitted the literature on logical representation of ambiguity, underspecification etc, except for the collection edited by van Deemter and Peters (1995). One strand of this research concerns what Apresjan called *regular polysemy* (such as the animal/meat uses of words such as *rabbit*, *halibut*, *chicken*, *elk*) and how to represent it, and the extent to which the meanings of such uses are predictable and the phenomenon of blocking (*beef* usually blocks the meat use of *cow* etc). Briscoe and Copestake (1991), Copestake (1995), Copestake and Briscoe (1995), Nunberg and Zaenen (1992), Ostler and Atkins (1992) discuss regular polysemy and its computational representation. Briscoe *et al.*(1995) provide a formal account of some aspects of blocking. Copestake and Briscoe (1995) argue that this sort of *sense extension* should be distinguished from *constructional polysemy*, where there is no lexical ambiguity, but where variations in meaning arise from the combination of lexical items. From this viewpoint, *logical metonymy* and *co-composition*, which are discussed by Pustejovsky and other authors working within the Generative Lexicon approach, are best seen as cases of constructional polysemy. Godard and Jayez (1993) describe some problems for Pustejovsky's account of logical metonymy as a coercion operation.

Clark and Clark (1979) and Clark (1992) discuss nonce uses of nouns as verbs and argue that meaning is not predictable and must be regarded as arising from the context. Nunberg (1978) discusses a range of problems which also point to some contribution of pragmatics in reference.

Schütze's (1995) model of ambiguity is quite different from most of the models discussed here, as it is based on the use of gradient rather than discrete representations. This appears to have advantages, at least when modelling language acquisition, although it does not allow generalisations to be expressed in the same way as the lexical rule models.

Apresjan, Ju D. (1973) *Regular polysemy,* Mouton, The Hague, The Netherlands.

Briscoe, E. J., and A. Copestake (1991) 'Sense extensions as lexical rules', *Proceedings of the IJCAI Workshop on computational approaches to non-literal language,* Sydney, Australia, pp. 12–20.
`http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`

Briscoe, E.J., A. Copestake and A. Lascarides (1995) 'Blocking' in P. Saint-Dizier, P. and E. Viegas (ed.), *Computational lexical semantics,* Cambridge University Press, Cambridge, UK.
`http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`

Clark, H. H. (1992) *Arenas of language use,* University of Chicago Press, Chicago and CSLI, Stanford.

Clark, E. V. and H. H. Clark (1979) 'When nouns surface as verbs', *Language,* **55**, 767–811.

Copestake, A (1995) 'Representing lexical polysemy', *Proceedings of the AAAI Spring Symposium on lexical semantics*, Stanford, CA.
`http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`

Copestake, A. and E.J. Briscoe (1995) 'Semi-productive polysemy and sense extension', *Journal of Semantics,* **12**, 15–67.
`http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`

Cruse, D. A. (1995) 'Polysemy and related phenomena from a cognitive linguistic viewpoint' in P. Saint-Dizier and E. Viegas (ed.), *Computational lexical semantics,* Cambridge University Press, Cambridge, UK, pp. 33–49.

van Deemter, K. and S. Peters (eds) (1995) *Semantic ambiguity and underspecification,* CSLI Publications.

Fauconnier, G. (1985) *Mental spaces: aspects of meaning construction in natural language,* MIT Press, Cambridge, MA.

Godard, D. and J. Jayez (1993) 'Towards a proper treatment of coercion phenomena', *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics (EACL-93),* Utrecht, The Netherlands, pp. 168–177.

Jackendoff, R. (1975) 'Morphological and semantic regularities in the lexicon', *Language,* **51(3)**, 639–71.

Kilgarriff, A. (1992) 'Polysemy', Doctoral dissertation, University of Sussex, Cognitive Science Research Paper CSRP 261.

Nunberg, G.D. (1995) 'Transfers of meaning', *Journal of Semantics,* **12**.

Nunberg, G.D. (1978) 'The pragmatics of reference', Doctoral Dissertation, CUNY Graduate Center, reproduced by the Indiana University Linguistics Club.

Nunberg, G.D. (1979) 'The non-uniqueness of semantic solutions: polysemy', *Linguistics and Philosophy,* **3**, 145–184.

Nunberg, G.D. (1993) 'On the meaning and interpretation of indexical expressions', *Linguistics and Philosophy,* **16**, 1–43.

Nunberg, G.D. and A. Zaenen (1992) 'Systematic polysemy in lexicology and lexicography', Proceedings of Euralex92, Tampere, Finland.

Ostler, N. and B.T. Atkins (1992) 'Predictable meaning shift: some linguistic properties of lexical implication rules' in J. Pustejovsky and S. Bergler (ed.), *Lexical semantics and knowledge representation. Proceedings of the first SIGLEX Workshop, Berkeley, CA,* Springer-Verlag, Berlin, pp. 87–100.

Pustejovsky, J. (1993) 'Type coercion and lexical selection' in J. Pustejovsky (ed.), *Semantics and the lexicon,* Kluwer, Dordrecht, pp. 73–96.

Pustejovsky, J. (1995) 'Linguistic constraints on type coercion' in P. Saint-Dizier and E. Viegas (ed.), *Computational lexical semantics,* Cambridge University Press, Cambridge, UK.

Schütze, H. (1995) 'Ambiguity in language learning: computational and cognitive models', PhD thesis, Stanford University
`ftp://parcftp.xerox.com/pub/qca/pubs.html`.

Zwicky, A. M. and J. M. Sadock (1975) 'Ambiguity tests and how to fail them' in J. P. Kimball (ed.), *Syntax and Semantics IV,* Academic Press, New York, pp. 1–36.

7 Metaphor, Idioms and Collocations

These issues are perhaps peripheral to the main body of work on lexical semantics, but nevertheless can throw some light on issues such as the relationship between pragmatics and lexical semantics. Metaphor is interesting in that there seem to be conventional patterns of metaphor (Lakoff, 1987; Lakoff and Johnson, 1980): some computational linguistic approaches have tried to capture this idea (Martin 1990). Briscoe and Copestake (1991) argue that, to the extent some metaphor patterns (such as animal names used for humans) are conventionalized, they may be partially represented via lexical rules. Asher and Lascarides (1995) try to link this idea with a model of discourse, in order to explain the meanings of metaphors in context.

I have included some papers on idioms which have concentrated on viewing them as a semantic phenomenon: Nunberg *et al.*(1994) being the classic account of this (see Pulman, 1993, for an alternative). Fellbaum (1993) argues that the behaviour of idioms with respect to determiners can be (partially) accounted for on semantic grounds.

The term *collocation* is used in the corpus processing literature to indicate any grouping of words which occur together more frequently than would be expected if the distribution were random. In terms of lexical semantics, it is useful to restrict its use to word groups which are essentially unpredictable. For example, the fact that *drive* and *car* co-occur relatively frequently is something that might be predicted from the meanings of these terms and real-world knowledge. In contrast, the collocation of *roast* and *beef* as opposed to *bake* and *beef* might not be predictable since it is difficult to supply a meaning for *roast* as opposed to *bake* that predicts that chestnuts, coffee beans and beef are roasted, while apples, fish and ham are baked. The assumption is that collocational restrictions are not part of meaning, and may be to some extent idiosyncratic, in which case they are presumably simply learned with respect to individual words (see Cruse, 1986:279).

This definition makes the idea of collocation dependent on precisely what the meaning of a term is taken to be. For example, if *rancid* is a general term for badness of food, then the fact that *rancid butter* is more usual than *rancid strawberries* would have to be explained as a collocation, whereas if *rancid* applies only to the particular way in which fats or oils characteristically become unpleasant tasting, the contrast is explicable using world knowledge about the composition of butter and strawberries.[2] According to this view, collocations differ from idioms in that the meaning of the phrase is transparent, though the term collocation is sometimes used to subsume both idioms and transparent but conventional phrases.

To the extent that computational theories of meaning abstract away from individual words, they have to allow for collocations, at least to support generation. But this isn't a topic that has received much attention in the computational linguistic literature. Mel'čuk's work however does concentrate heavily on collocations, though the term is used in a somewhat different way. See Heylen *et al.*(1993) for an account of the contrasting notions of collocations and an implementation of the Mel'čukian notion in a constraint-based framework.

Asher, N. and A. Lascarides (1995) 'Metaphor in discourse', *Proceedings of the AAAI Spring Symposium on Lexical Semantics*, Stanford CA.

Briscoe, E. J., and A. Copestake (1991) 'Sense extensions as lexical rules', *Proceedings of the IJCAI Workshop on Computational Approaches to Non-Literal Language,* Sydney, Australia, pp. 12–20.
`http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`

Fass, D. (1991) 'met*: a method for discriminating metonymy and metaphor by computer', *Computational Linguistics,* **17:1**, 49–90.

Fellbaum, C. (1993) 'The determiner in English idioms' in C. Cacciari and P. Tabossi (ed.), *Idioms: processing, structure, and interpretation,* Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 271–296.

Heylen, D., A. Schenk, M. Verhagen (1993) 'A constraint-based representation scheme of collocational structures' in W. Sijtsma and O. Zweekhorst (ed.), *3rd Computational Linguistics in the Netherlands (CLIN) Meeting,* ITK, Tilburg, pp. 49–58.
`http://wwwots.let.ruu.nl/Publications/occasions.html`

[2] It's perfectly possible that one man's meaning is another man's collocation.

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

Lakoff, G. and M. Johnson (1980) *Metaphors we live by,* University of Chicago Press, Chicago.

Martin, J. (1990) *A computational model of metaphor interpretation,* Academic Press, Cambridge, MA.

Mel'čuk, I. and M. Polguère (1987) 'A formal lexicon in Meaning-Text Theory (or how to do lexica with words)', *Computational Linguistics,* **13:3–4**, 261–275.

Nunberg, G., T. Wasow and I. Sag (1994) 'Idioms', *Language,* **70**, 491–538.

Pulman, S.G. (1993) 'The recognition and interpretation of idioms' in C. Cacciari and P. Tabossi (ed.), *Idioms: processing, structure, and interpretation,* Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 249–270.

Sweetser, E. (1990) *From etymology to pragmatics,* Cambridge University Press, Cambridge, UK.

8 Formal representation: inheritance and defaults

I have only listed papers which are based on some sort of feature structure framework, but Daelmans *et al.*(1992) is an overview of inheritance in NLP generally. Shieber (1986), although somewhat dated, is still the most accessible introduction to unification-based grammars. DATR (Evans and Gazdar, 1989a,b) is a formalism designed for lexical representation which can be interfaced with a unification-based system for parsing etc. Other work has avoided making such a strong distinction between the lexicon and other system components. Carpenter (1992) is a very detailed coverage of one approach to typed feature structures. I have not listed any of the other literature on typed feature structures here, since none of it is directly relevant to lexical semantics, (though see the bibliography of the short paper *Inheritance in lexical representation*, accessible via `http://www-csli.stanford.edu/users/aac`). Flickinger's work concentrates on syntax but introduced the idea of the hierarchical lexicon as a major theme within HPSG. Copestake *et al.*(1993), Sanfilippo (1993) and other papers in Briscoe *et al.* (1993) describe how typed feature structure formalisms may be used in lexical representation: Copestake (1993a) gives the details of the implemented system assumed there. Krieger and Nerbonne (1993) discuss inheritance and lexical rules for the representation of morphology (see also Riehemann, 1994).

In contrast with DATR and Flickinger's work, most typed feature structure systems assume strict monotonicity of inheritance, but there have been a number of papers on priority union or default unification (van den Berg and Prüst (1991); Bouma, 1990, 1992; Calder, 1991; Carpenter, 1993; Copestake, 1993b; Kaplan, 1987; Lascarides *et al.*1996; Lascarides and Copestake, 1995 (this was actually written after the preceding paper and suggests a modification of the approach proposed there); Russell *et al.*, 1991, 1993; Young and Rounds, 1993). Most of these discuss the operation in rather abstract terms, or with reference to morphology, but default inheritance is an important representation technique in lexical semantics, since so many of the generalizations that one would like to express have exceptions. Copestake (1992) and Lascarides *et al.*(1996) give details of the use of defaults in lexical semantics. Kilgarriff (1993) shows how some lexical semantic information may be encoded in DATR.

van den Berg, M. and H. Prüst (1991) 'Common denominators and default unification', *Proceedings of the Computational Linguistics in the Netherlands (CLIN), First Meeting,* Utrecht, pp. 1–16.

Bouma, G. (1990) 'Defaults in unification grammar', *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics (ACL-90),* Pittsburg, Pennsylvania, pp. 165–173.

Bouma, G. (1992) 'Feature structures and nonmonotonicity', *Computational Linguistics,* **18.2**, 183–204.

Briscoe, E. J., A. Copestake and V. de Paiva (eds) (1993) *Inheritance, defaults and the lexicon,* Cambridge University Press, Cambridge, UK.

Calder, J. (1991) 'Some notes on priority union', Paper presented at the ACQUILEX Workshop on Default Inheritance in the Lexicon, Cambridge.

Carpenter, R. (1992) *The logic of typed feature structures,* Cambridge University Press, Cambridge, UK.

Carpenter, R. (1993) 'Skeptical and credulous default unification with application to templates and inheritance' in E. J. Briscoe, A. Copestake and V. de Paiva (ed.), *Inheritance, defaults and the lexicon,* Cambridge University Press, Cambridge, UK, pp. 13–37.

Copestake, A. (1992) 'The representation of lexical semantic information', Doctoral dissertation, University of Sussex, Cognitive Science Research Paper CSRP 280 `http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`.

Copestake, A. (1993a) 'The Compleat LKB', ACQUILEX-II Deliverable, 3.1 `http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`.

Copestake, A. (1993b) 'Defaults in Lexical Representation' in E. J. Briscoe, A. Copestake and V. de Paiva (ed.), *Inheritance, defaults and the lexicon,* Cambridge University Press, pp. 223–245.

Copestake, A., A. Sanfilippo, E.J. Briscoe and V. de Paiva (1993) 'The ACQUILEX LKB: an introduction' in E. J. Briscoe, A. Copestake and V. de Paiva (ed.), *Inheritance, defaults and the lexicon,* Cambridge University Press, Cambridge, UK, pp. 148–163.

Daelemans, W., K. de Smedt and G. Gazdar (1992) 'Inheritance in natural language processing', *Computational Linguistics,* **18.2**, 205–218.

Evans, R. and G. Gazdar (1989a) 'Inference in DATR', *Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics (EACL-1989),* Manchester, UK, pp. 66–71.

Evans, R. and G. Gazdar (1989b) 'The semantics of DATR' in A. G. Cohn (ed.), *Proceedings of the Seventh Conference of the Society for the Study of Artificial Intelligence and Simulation of Behavior (AISB),* Pitman/Morgan Kaufmann, London, pp. 79–87.

Flickinger, D. (1987) 'Lexical rules in the hierarchical lexicon', PhD thesis, Stanford University `http://hpsg.stanford.edu/hpsg/papers.html`.

Flickinger, D., C. Pollard and T. Wasow (1985) 'Structure sharing in lexical representation', *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics (ACL-85),* University of Chicago, pp. 262–268.

Flickinger, D. and J. Nerbonne (1992) 'Inheritance and complementation: a case study of *easy* adjectives and related nouns', *Computational Linguistics,* **18.3**, 269–310.

Kaplan, R. (1987) 'Three seductions of computational psycholinguistics' in P. Whitelock, H. Somers, P. Bennett, R. Johnson, and M. McGee Wood (ed.), *Linguistic theory and computer applications,* Academic Press, London, pp. 149–188.

Kilgarriff, A. (1993) 'Inheriting verb alternations', *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics (EACL-93),* Utrecht, The Netherlands, pp. 213–221.

Krieger, H-U. and J. Nerbonne (1993) 'Feature-based inheritance networks for computational lexicons' in E. J. Briscoe, A. Copestake and V. de Paiva (ed.), *Inheritance, defaults and the lexicon,* Cambridge University Press, Cambridge, UK, pp. 90–137.

Lascarides, A., E.J. Briscoe, N. Asher and A. Copestake (1996) 'Persistent associative default unification', *Linguistics and Philosophy,* **19:1**, 1–89. `http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`

Lascarides, A. and A. Copestake (1995) 'A logic for order-independent typed default unification', Acquilex-II Working paper 60
http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html.

Riehemann, S. (1994) 'Morphology and the hierarchical lexicon', Stanford University
http://hpsg.stanford.edu/hpsg/papers.html.

Russell, G., J. Carroll and S. Warwick (1991) 'Multiple default inheritance in a unification-based lexicon', *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics (ACL-91),* Berkeley, California, pp. 215–221.

Russell, G., A. Ballim, J. Carroll and S. Warwick-Armstrong (1993) 'A practical approach to multiple default inheritance for unification-based lexicons' in E. J. Briscoe, A. Copestake and V. de Paiva (ed.), *Inheritance, defaults and the lexicon,* Cambridge University Press, Cambridge, UK, pp. 137–147.

Sanfilippo, A. (1993) 'LKB encoding of lexical knowledge' in E. J. Briscoe, A. Copestake and V. de Paiva (ed.), *Inheritance, defaults and the lexicon,* Cambridge University Press, Cambridge, UK, pp. 190–222.

Shieber, S. M. (1986) *An introduction to unification-based approaches to grammar,* CSLI Lecture Notes 4, Stanford CA.

Young, M. and W. Rounds (1993) 'A logical semantics for non-monotonic sorts', *Proceedings of the 31st Conference of the Association for Computational Linguistics (ACL-93),* Columbus, Ohio, pp. 209–215.

9 Formal representation: lexical rules

The formalization of lexical rules in typed feature structure frameworks and other similar approaches is somewhat problematic: some of the papers listed here are an attempt to address this (see also Riehemann, 1994, listed above). Most of the papers discussing the use of lexical rules in lexical semantics are listed in the polysemy section.

Briscoe, E.J. and A. Copestake (1995) 'Dative constructions as Lexical rules in the TDFS framework', Acquilex-II Working Papers 78
(revised version should be available soon via http://www-csli.stanford.edu/users/aac.

Calcagno, M. (1995) 'Interpreting lexical rules', Proc. of the Conference on Formal Grammar, Barcelona.

Flickinger, D. (1987) 'Lexical rules in the hierarchical lexicon', PhD thesis, Stanford University
http://hpsg.stanford.edu/hpsg/papers.html.

Meurers, D. (1995) 'Towards a semantics for lexical rules as used in HPSG', Acquilex Workshop on Lexical Rules, Cambridge
http://www.sfs.nphil.uni-tuebingen.de/~dm/.

Meurers, D. and G. Minnen (1995) 'A computational treatment of HPSG lexical rules as covariation lexical entries', *Proceedings of the 5th Int. Workshop on Natural Language Understanding and Logic Programming,* Lisbon.
http://www.sfs.nphil.uni-tuebingen.de/~dm/

10 Disambiguation

For NLP the issues of polysemy and disambiguation are closely intertwined, since if lexical ambiguity is postulated, it (usually) has to be resolved. While it is clear that arbitrary word knowledge may be necessary to resolve ambiguity, and therefore that ambiguity resolution is not a good criterion for deciding whether or not knowledge should be regarded as lexical (for systems that distinguish between lexical and non-lexical semantics), high success rates on natural text can be achieved with cruder information. McRoy (1992) is a reasonable source for references to work on disambiguation prior to 1990, but recent work has mostly concentrated on the use of statistical techniques.

Asher, N. and A. Lascarides (1995) 'Lexical disambiguation in a discourse context', *Journal of Semantics,* **12(1)**, 69–108.

Bruce, R. and J. Wiebe (1994) 'Word-sense disambiguation using decomposable models', *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics (ACL-94),* Las Cruces, NM, pp. 139–146.

Dagan, I. and A. Itai (1994) 'Word-sense disambiguation using a second language monolingual corpus', *Computational Linguistics,* **20(4)**, 563–596.

Gale, W., K. Church and D. Yarowsky (1992) 'A method for disambiguating word senses in a large corpus', *Computers and the Humanities,* **26**, 415–439.

Harley, A. (1994) 'Cambridge Language Survey: semantic tagger', ACQUILEX-II working paper No 39: available from `cide@cup.cam.ac.uk`.

Justeson, J.S. and S.M. Katz (1995) 'Principled disambiguation: discriminating adjective senses with modified nouns', *Computational Linguistics,* **21(1)**, 1–28.

McRoy, S. (1992) 'Using multiple knowledge sources for word sense discrimination', *Computational Linguistics,* **18(1)**, 1–30.

Resnik, P. (1995) 'Disambiguating noun groupings with respect to WordNet senses', *Proceedings of the 3rd workshop on very large corpora,* MIT, Cambridge, Mass..
`http://xxx.lanl.gov/cmp-lg/9511006`

Schütze, H. (1995) 'Ambiguity in language learning: computational and cognitive models', PhD thesis, Stanford University
`ftp://parcftp.xerox.com/pub/qca/pubs.html`.

Yarowsky, D. (1992) 'Word-sense disambiguation using statistical models of Roget's categories trained on large corpora', *Proceedings of the 14th International Conference on Computational Linguistics (Coling 92),* Nantes, France.

11 Acquision of lexical semantics from corpora and machine readable dictionaries

One of the main problems in assuming detailed lexical semantic information in a lexicon has been the issue of how to acquire it. Until recently the dominant technique involved the use of machine-readable dictionaries (MRDs). I haven't attempted to summarize this literature here. Boguraev and Briscoe (1989) contains some of the most interesting work and has a good introduction, Briscoe (1991) and Calzolari (1992) provide a partial summary of some subsequent work. However I've tried to be reasonably comprehensive in the references to recent work, at least with respect to papers published in the main conferences, since the MRD literature is often quite bad with respect to citations. Unfortunately there is a lot of work that is not easy to find: in particular, a lot of Vossen's work is difficult to get (try emailing `Piek.Vossen@let.uva.nl`). Zernik (1992) contains papers on both corpus processing and the use of MRDs. Since that collection was published, the use of corpora has increased, although there is still more work on extraction of syntactic information than on lexical semantics. A major problem with the assessment of this work is that there is very little agreement on what lexical semantic information is required.

Atkins, B.T. and B. Levin (1992) 'Admitting impediments' in U. Zernik (ed.), *Lexical acquisition: using on-line resources to build a lexicon,* Lawrence Erlbaum, New Jersey.

Basili, R., M. Pazienza, and P. Velardi (1992) 'Computational lexicons: the neat examples and the odd exemplars', *Proceedings of the 3rd Conf. on Applied Natural Language Processing,* Trento, Italy, pp. 96–103.

Boguraev, B. and E.J. Briscoe (eds) (1989) *Computational lexicography for natural language processing,* Longman, London.

Briscoe, E.J. (1991) 'Lexical issues in natural language processing' in E. Klein and F. Veltman (ed.), *Natural language and speech,* Springer-Verlag, pp. 39–68.

Calzolari, N. (1992) 'Acquiring and representing semantic information in a lexical knowledge base' in J. Pustejovsky and S. Bergler (ed.), *Lexical semantics and knowledge representation,* Lecture Notes in AI: 627, Springer-Verlag, Berlin, pp. 235–244.

Copestake, A. (1990) 'An approach to building the hierarchical element of a lexical knowledge base from a machine readable dictionary', *Proceedings of the First International Workshop on Inheritance in Natural Language Processing,* Tilburg, The Netherlands, pp. 19–29.
`http://www.cl.cam.ac.uk/Research/NL/acquilex/papers.html`

Justeson, J.S. and S.M. Katz (1991) 'Co-occurrences of antonymous adjectives and their contexts', *Computational Linguistics,* **17(1)**, 1–20.

Klavans, J.L. (1988) 'COMPLEX: a computational lexicon for natural language systems', *Proceedings of the 12th International Conference on Computational Linguistics (Coling-88),* Budapest, Hungary.

Klavans, J.L., M.S. Chodorow and N. Wacholder (1990) 'From dictionary to knowledge base via taxonomy', *Proceedings of the Sixth Conference of the University of Waterloo Centre for the New Oxford English Dictionary and Text Research: Electronic Text Research,* University of Waterloo, Canada.

Light, M. (1996) 'Morphological cues for lexical semantics', *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics (ACL-96),* Santa Cruz, CA.
`http://www.sfs.nphil.uni-tuebingen.de/~light/morphcues.ps.gz`

Montemagni S. and L. Vanderwende (1992) 'String Patterns versus Structural Patterns for Extracting Semantic Information from Dictionaries', *Proceedings of the 14th International Conference on Computational Linguistics (Coling 92),* Nantes, France.

Pustejovsky, J., S. Bergler, P. Anick (1993) 'Lexical semantic techniques for corpus analysis', *Computational Linguistics,* **19(2)**, 331–358.

Sanfilippo, A. and V. Poznanski (1992) 'The acquisition of lexical knowledge from combined machine-readable dictionary sources', *Proceedings of the 3rd Conference on Applied Natural Language Processing (ANLP-92),* Trento, Italy, pp. 80–88.

Velardi, P., M. Pazienza, and M. Fasolo (1991) 'How to encode semantic knowledge: a method for meaning representation and computer-aided acquisition', *Computational Linguistics,* **17:2**, 153–170.

Vossen, P. (1990) 'A parser-grammar for the meaning descriptions of LDOCE', Links Project Technical Report 300-169-007, Amsterdam University.

Vossen, P. (1992) 'The automatic construction of a knowledge base from dictionaries: a combination of techniques' in H. Tommola, K. Tarantola, T. Salmin Tolonen, J. Schop (ed.), *Euralex '92 Proceedings I-II,* Tampere, Finland.

Vossen, P. (1995) *Grammatical and conceptual individuation in the lexicon,* IFOTT, Amsterdam.

Vossen, P., W. Meijs and M. Broeder (1989) 'Meaning and structure in dictionary definitions' in B. Boguraev and E.J. Briscoe (ed.), *Computational lexicography for natural language processing,* Longman, London, pp. 171–192.

Vossen, P. and A. Copestake (1993) 'Untangling definition structure into knowledge representation' in E. J. Briscoe, A. Copestake and V. de Paiva (ed.), *Inheritance, defaults and the lexicon,* Cambridge University Press, Cambridge, UK, pp. 246–274.

Walker, D. and R. Amsler (1986) 'The use of machine-readable dictionaries in sublanguage analysis' in R. Grishman and R. Kittredge (ed.), *Analyzing language in restricted domains,* Lawrence Erlbaum Associates, Hillsdale, New Jersey, pp. 69–83.

Wilks, Y., D. Fass, C-M. Guo, J. McDonald, T. Plate and B. Slator (1989) 'A tractable machine dictionary as a resource for computational semantics' in B. Boguraev and E.J. Briscoe (ed.), *Computational lexicography for natural language processing,* Longman, London, pp. 193–231.

Zernik, U. (ed.) (1992) *Lexical acquisition: using on-line resources to build a lexicon,* Lawrence Erlbaum, New Jersey.