

Chapter 10

WordNet

Christiane Fellbaum

10.1 Introduction

WordNet is a large electronic lexical database for English (Miller, 1995; Fellbaum, 1998a). It originated in 1986 at Princeton University where it continues to be developed and maintained. George A. Miller, a psycholinguist, was inspired by experiments in Artificial Intelligence that tried to understand human semantic memory (e.g., Collins and Quillian, 1969). Given the fact that speakers possess knowledge about tens of thousands of words and the concepts expressed by these words, it seemed reasonable to assume efficient and economic storage and access mechanisms for words and concepts. The Collins and Quillian model proposed a hierarchical structure of concepts, where more specific concepts inherit information from their superordinate, more general concepts; only knowledge particular to more specific concepts needs to be stored with such concepts. Thus, it took subjects longer to confirm a statement like “canaries have feathers” than the statement “birds have feathers” since, presumably, the property “has-feathers” is stored with the concept *bird* and not redundantly with the concept for each kind of bird.

While such theories seemed to be confirmed by experimental evidence based on a limited number of concepts only, Miller and his team were asking whether the bulk of the lexicalized concepts of a language could be represented with hierarchical relations in a network-like structure. The result was WordNet, a large, manually constructed semantic network where words that are similar in meaning are inter-related. While WordNet no longer aims to model human semantic organization, it has become a major tool for Natural Language Processing and spawned research in lexical semantics and ontology.¹

C. Fellbaum (✉)

Department of Computer Science, Princeton University, Princeton, NJ 08540, USA
e-mail: fellbaum@princeton.edu

¹For critical reviews of WordNet, see Kilgarriff (2000) and Lin (1999).

10.2 Design and Contents

WordNet is a large semantic network – a graph – in which words are interconnected by means of labeled arcs that represent meaning relations. Lexical relations connect single words while semantic-conceptual relations links concepts that may be expressed by more than one word.

Synonymy is the many-to-one mapping of word forms and concepts. For example, both the strings *boot* and *trunk* can refer to the same concept (the luggage compartment of a car). Under these readings, the two word forms are synonyms. WordNet groups synonyms into unordered sets, called synsets. Substitution of a synset member by another does not change the truth value of the context, though one synonym may be stylistically more felicitous than another in some contexts.

A synset lexically expresses a concept. Examples of synsets – marked here by curly brackets – are mail, post, hit, strike, and small, little. WordNet’s synsets further contain a brief definition, or “gloss,” paraphrasing the meaning of the synset, and most synsets include one or more sentences illustrating the synonyms’ usage. A domain label (“sports,” “medicine,” “biology,” etc.) marks many synsets.

Polysemy is the many-to-one mapping of meanings to word forms. Thus, *trunk* may refer to a part of a car, a tree trunk, a torso, or an elephant’s proboscis. In WordNet, membership of a word in multiple synsets reflects that word’s polysemy, or multiplicity of meaning. *Trunk* therefore appears in several different synsets, each with its own synonyms. Similarly, the polysemous word form *boot* appear in several synsets, once together with *trunk*, another time as a synonym of *iron boot* and *iron heel*, etc.

Synsets are the nodes or building blocks of WordNet. As a result of the interconnection of synsets via meaning-based relations, a network structure arises.

10.3 Coverage

WordNet in fact consists of four separate parts, each containing synsets with words from the major syntactic categories: nouns, verbs, adjectives, and adverbs. The current version of WordNet (3.0) contains over 117,000 synsets, comprising over 81,000 noun synsets, 13,600 verb synsets, 19,000 adjective synsets, and 3,600 adverb synsets. The separation of words and synsets for different parts of speech follows from the nature of the word class-specific semantic and lexical relations.

10.4 Relations

Besides synonymy, WordNet encodes another lexical (word-word) relation, antonymy (or, more generally, semantic contrast or opposition). Antonymy is psychologically salient, particularly among adjective pairs like *wet-dry* and *long-short*, but it is also encoded for verb pairs like *rise-fall* and *come-go*. (WordNet does not make the kind of subtle distinctions among the different kinds of semantic opposition drawn in, e.g., Cruse (1986)).

Another kind of lexical relation, dubbed “morphosemantic,” is the only one that links words from all four parts of speech. It connects words that are both morphologically and semantically related (Fellbaum and Miller, 2003). For example, the semantically related senses of *interrogation*, *interrogator*, *interrogate*, and *interrogative* are interlinked.

All other relations in WordNet are conceptual-semantic relations and connect not just words (synset members) but entire synsets. For each part of speech, different relations were identified.

10.5 Nouns in WordNet

Nouns comprise the bulk of the English lexicon, and the noun component of WordNet reflects this. Nouns are relatively easy to organize into a semantic network; WordNet largely follows the Aristotelian model of categorization.

10.5.1 Hyponymy

Noun synsets are primarily interconnected by the hyponymy relation (or hyperonymy, or subsumption, or the ISA relation), which links specific concepts to more general ones. For example, the synset gym shoe, sneaker, tennis shoe is a hyponym, or subordinate, of shoe, which in turn is a hyponym of footwear, footgear, etc. And gym shoe, sneaker, tennis shoe is a hypernym, or superordinate, of plimsoll, which denotes a specific type of sneaker. The relation is bi-directional; therefore, these examples express both that gym shoe, sneaker, tennis shoe is a type of footwear, footgear and that the category footwear, footgear is comprised of gym shoe, sneaker, tennis shoe (as well as other types of footwear, such as boot and over-shoe). Hyponymy is transitive, so plimsoll, by virtue of being a type of gym shoe, sneaker, tennis shoe, which is a type of footwear, footgear, is also a type of footwear, footgear.

Hyponymy builds hierarchical “trees” with increasingly specific “leaf” concepts growing from an abstract “root.” Noun hierarchies can be deep and comprise as many as fifteen layers, particularly for biological categories, where WordNet includes both expert and folk terms.

All noun synsets ultimately descend from a single root, *entity*. The next layer comprises three synsets: *physical entity*, *abstract entity*, and *thing*. Below these, we find the synsets *object*, *living thing*, *causal agent*, *matter*, *physical process*, *substance*, *psychological feature*, *attribute*, *group*, *relation*, *communication*, *measure*, *quantity*, *amount*, and *otherworld*.

The selection of these very broad categories was of course somewhat subjective and has engendered discussion with ontologists. On an empirical level, it remains to be seen whether wordnets for other languages draw the same fundamental distinctions.²

²We will refer to the Princeton WordNet as “WordNet” and databases in other languages as “wordnets,” indicating the fact that the proper name *WordNet* has become a common noun.

10.5.2 *Types vs. Instances*

Within the noun hierarchies, WordNet distinguishes two kinds of hyponymys, types and instances. Common nouns are types: *city* is a type of *location*, and *university* is a type of *educational establishment*. However, *New York* and *Princeton* are not types, but instances of *city* and *educational establishment*, respectively. Proper names are instances, and instances are always leaf nodes that have no hyponyms (Miller and Hristea, 2004).

While the Princeton WordNet does not distinguish roles from types and instances, some later wordnets do, e.g., EuroWordNet (Vossen, 1998). Thus, nouns like *pet* and *laundry* are encoded as types of *animal* and *garment*, respectively, on par with *poodle* and *robe*. This treatment does not satisfactorily reflect the categorial status of such nouns; on the other hand, it is doubtful whether a consistent labeling of role nouns is possible (David Israel, personal communication).

10.5.3 *Meronymy*

Another major relation among noun synsets is meronymy (or part-whole relation). It links synsets denoting parts, components, or members to synsets denoting the whole. Thus, *toe* is a meronym of *foot*, which in turn is a meronym of *leg* and so on. Like hyponymy, meronymy is bi-directional. WordNet tells us that a *foot* has *toes* and that *toe* is a part of a *foot*. Hyponyms inherit the meronyms of their superordinates: If a *car* has *wheels*, then kinds of *cars* (*convertible*, *SUV*, etc.) also have *wheels*. (But note that statements like “a toenail is a part of a leg,” though true, sound odd.)

Meronymy in WordNet actually encompasses three semantically distinct part-whole relations. One holds among proper parts or components, such as *feather* and *wing*, which are parts of *bird*. Another links substances that are constituents of other substances: *oxygen* is a constituent part of *water* and *air*. Members like *tree* and *student* are parts of groups like *forest* and *class*, respectively. Many more subtle kinds of meronymy could be distinguished (Chaffin, 1992).

10.6 Verbs

Verbs are fundamentally different from nouns in that they encode events and states that involve participants (expressed by nouns) and in that they have temporal extensions. The classic Aristotelian relations that work well to construct a network of noun synsets are not suitable for connecting verbs. Verb synsets are organized by several lexical entailment relations (Fellbaum, 1998b). The most frequently encoded relation is “troponymy”, which relates synset pairs such that one expresses a particular manner of the other. For example, *mumble* is a troponym of *talk*, and *scribble* is a troponym of *write*. Like hyponymy, troponymy builds hierarchies with several levels of specificity, but verb hierarchies are more shallow than noun hierarchies and rarely exceed four levels.

The particular manner encoded by troponyms is not specified, and troponymy is in fact a highly polysemous relation whose semantics are domain-dependent. For communication verbs, the medium distinguishes broad classes of verb hierarchies (*speaking, writing, gesturing*); motion verbs tend to be differentiated by such components as speed (*walking vs. running vs. ambling*).

Another relation among verb synsets is backward entailment, where the event encoded in one synset necessarily entails a prior event that is expressed by the second synset. Examples are *divorce* and *marry* and *untie* and *tie*. While the events in such pairs do not temporally overlap, those linked via a presupposition relation do. Examples are *buy* and *pay*: If someone buys something, he necessarily pays for it, and paying is a necessary part of the buying event. Finally, WordNet encodes a cause relation, as between *show* and *see* and *raise* and *rise*. Note that these relations are unidirectional.

A particular kind of polysemy is found in “auto-relations,” where a word form has a sense that expresses both the general and the specific concept, as in *drink*, *imbibe* and *drink*, *booze* (Fellbaum, 2000).

10.7 Adjectives

Antonymy is the prevailing relation among adjectives. Most adjectives are organized into “direct” antonym pairs, such as *wet-dry* and *long-short*.

Each member of a direct antonym pair is associated with a number of “semantically similar” adjectives, either near-synonyms or different values of a given scalar property. Thus, *damp* and *drenched* are semantically similar to *wet*, while *arid* and *parched* are similar to *dry*. These semantically similar adjectives are said to be “indirect” antonyms of the direct antonym of their central members, i.e., *drenched* is an indirect antonym of *dry* and *arid* is an indirect antonym of *wet* (Miller, 1998). For experimental work examining this theory see Gross et al., (1989).

WordNet also contains “relational” adjectives, which are morphologically derived from, and linked to, nouns in WordNet. An example is *atomic*, *nuclear*, which is linked to *atom*, *nucleus*.

10.8 Where do Relations Come from?

People often ask how the WordNet relations and the specific encodings were arrived at. Some of the relations, like hyponymy and meronymy, have been known since Aristotle. They are also implicitly present in traditional lexicographic definitions; a noun is typically defined in terms of its superordinated and the particular differentiae, or in terms of the whole entity of which the noun denotes a part. Verbs, too, are often defined following the classical genus-differentiae form.

Word association norms compile the responses people give to a lexical stimulus. Frequent responses are words that denote subordinate and superordinate concepts, or words that are semantically opposed to the stimulus words.

For adjectives, the responses are strikingly uniform and robust; thus, most people say *cold* when asked to respond with the word that comes to mind first when they hear *hot*. These data inspired the organization of adjectives in terms of antonymy (Miller, 1998).

To encode these relations for specific words and synsets, the WordNet team relied on existing lexicographic sources as well as on introspection. In addition, Cruse (1986) lists some test for synonymy and hyponymy. For example, the pattern “Xs and other Ys” identifies X as a hyponym (subordinate) of Y, rather than a synonym.

When the bulk of WordNet was compiled, large corpora were not yet available that could have provided a different aspect on semantic similarity: co-occurrence in identical or similar contexts. More recent lexicons are often constructed semi-automatically, relying heavily on the distributional patterns of word forms as a measure of similarity.

10.9 WordNet as a Thesaurus

Traditional paper dictionaries are necessarily organized orthographically so as to enable look-up. But this means that words that are semantically related are not found together, and a user trying to understand the meanings of words in terms of related words or words in the definition of the target word, must flip many pages.

By contrast, WordNet’s semantics-based structure allows targeted look-up for meaning-related words and concepts from multiple access points. But unlike in a traditional thesaurus such as Roget’s, the arcs among WordNet’s words and synsets express a finite number of well-defined and labeled relations.

10.10 Semantic Distance and Lexical Gaps

The WordNet relations outlined here sufficed to interrelate the words of English; this was not at all obvious from the start. But WordNet’s apparently simple structure hides some unevenness. First, the meaning difference, or semantic distance, between parent and child nodes varies. For example, while verbs like *whisper*, *mumble*, and *shout* all seem equidistant from their parent *talk*, the distance between *talk* and its direct superordinate, *communicate*, seems much larger. This can be seen in the fact that *whisper*, *mumble* and *shout* can be fairly easily replaced by *talk* in many contexts without too much loss of information, whereas the substitution of *talk* with *communicate* would be very odd in many contexts.

A question related to semantic distance concerns lexical gaps, arguably concepts that for no principled reason are not linguistically labeled. For example, the lexicon suggests that nouns like *car*, *bicycle*, *bus*, and *sled* are all direct subordinates of *vehicle*. But this group of “children” seems heterogeneous: *sled* stands out for several reasons, in particular for not having wheels. To draw what appears like a

major distinction among the many vehicles, WordNet introduced a synset *wheeled vehicle*. The argument is that people distinguish between the category of wheeled vehicles and vehicles moving on runners independently of whether this distinction is lexically encoded in their language. One would expect other languages to label these concepts and show that the lack of an English word is purely accidental. (In fact, German has a word, *Kufenfahrzeug*, for *vehicle on runners*).³

Adding nodes in places where lexical gaps are perceived reduces the semantic distance among lexicalized categories but presents a more regular picture of the lexicalization patterns than warranted by purely linguistic data. Thus, the introduction of lexical gaps is a matter of discussion among wordnet builders. On the other hand, it is common practice in ontology, where it is usually assumed concepts are independent of natural language labels.

10.11 WordNet as an Ontology

Because of its rigid structure, WordNet is often referred to as an ontology; indeed, some philosophers working on ontology have examined WordNet's upper structure and commented on it. For example, Gangemi et al. (2002a, b) and Oltramari et al. (2002) have made specific suggestions for making WordNet more consistent with ontological principles. But the creators of WordNet prefer to call it a "lexical ontology," because its contents – with few exceptions – are concepts that are linguistically encoded and its structure is largely driven by the the lexicon. By contrast, many ontologists emphasize that an ontology is language-independent and merely uses language to refer to concepts and relations. Ontologies are usually understood to be knowledge structures rather than lexicons. For further discussion on the lexicon-ontology difference see Pease and Fellbaum (2009).

10.12 WordNet and Formal Ontology

WordNet has been linked to formal ontologies (Gangemi et al., 2002a; Niles and Pease, 2003). Concepts in one ontology, SUMO (Suggested Upper Merged Ontology, Niles and Pease, 2001; Niles and Pease, 2003; Chapter 11, Controlled English to Logic Translation, Pease and Li, this volume) have been linked to synsets not only in the Princeton WordNet but to many wordnets in other languages as well.

SUMO is a formal ontology stated in a first-order logic language called SUO-KIF. SUMO contains some 1,000 terms and 4,000 axioms using those terms in SUO-KIF statements. These axioms include some 750 rules. SUMO is an upper ontology, covering very general notions in common-sense reality, such as time, spatial relations, physical objects, events and processes.

³Fellbaum and Kegl (1989) argue for lexical gaps in the verb lexicon on syntactic grounds.

A mid-level ontology (MILO) was created to extend SUMO with several thousand more terms and associated definitions for concepts that are more specific. In addition, domain ontologies cover over a dozen areas including world government, finance and economics, and biological viruses. Together with SUMO and MILO they include some 20,000 terms and 60,000 axioms.

Niles and Pease (2003) manually mapped the formally defined terms in SUMO to synsets in WordNet. Three types of mappings were made: rough equivalence, subsuming, and instance. In addition, mappings were made for senses that appeared to occur frequently in language use, based on the SemCor semantic concordance (Miller et al., 1993). New concepts were created in the MILO as needed and linked to the appropriate synsets. SUMO, MILO, and the domain ontologies have been linked to wordnets in several other languages as well (for details on the linking see Pease and Fellbaum, 2009).

For example, the synset *artificial satellite, orbiter, satellite* maps to the formally defined term of *ArtificialSatellite* in SUMO. The mapping is an “equivalence” mapping since there is nothing that appears to differentiate the linguistic notion from the formal term in this case. A more common case of mapping is a “subsuming” mapping. For example *elk* maps to the SUMO term *HoofedMammal*. WordNet is considerably larger than SUMO and so many synsets map to the same more general formal term. As an example of an “instance” link, the synset *george washington, president washington, washington* is linked to the SUMO term *Human*. Because WordNet discriminates among different senses of the same linguistic token, the synset *evergreen state, wa, washington* is linked via an “instance” relation to the term *StateOrProvince*.

10.13 Wordnets in Other Languages

Since the 1990s, wordnets are being built in other languages. The first, EuroWordNet (EWN, Vossen, 1998), encompasses eight languages, including non-Indo-European languages like Estonian and Turkish. EuroWordNet introduced some fundamental design changes that have been adopted by many subsequent wordnets. Crucially, all wordnets are linked to the Princeton WordNet.

10.14 The EuroWordNet Model

Wordnets were constructed for each language following one of two strategies. The first, dubbed “Expand,” was to translate the synsets of the Princeton WordNet into the target language, making adjustments as needed (see below). The second, dubbed “Merge,” was to develop a semantic network in the target language from scratch and subsequently link it to the Princeton WordNet.

Several innovations were introduced. In contrast to Princeton WordNet’s strict limitation to paradigmatic relations, the wordnets built for EWN encode many cross-POS links. For example, syntagmatically associated nouns and verbs, such as the

pair *student* and *learn* are linked. Another innovation are conjunctive and disjunctive relations. Conjunctive relations allow a synset to have multiple superordinates. Thus *knife* is both a kind of *eating utensil* and a kind of *weapon*. Another example is *albino*, which can be a kind of *person*, *animal*, or *plant*. Double parenthood can capture the Type vs. Role distinction discussed earlier.

An example for a disjunctive relation is *airplane* and its meronyms *propellor* and *jet*; a given type of airplane has either, but not both, parts (Vossen, 1998). The possibility of disjoint parts can reduce the proliferation of artificial nodes, such as *propeller plane* and *jet plane*.

Each wordnet that is part of EuroWordNet relates to three language-neutral components: the Top Concept Ontology, the Domain Ontology, and Interlingual Lexical Index (ILI).

The Top Concept Ontology is a hierarchically organized set of some 1,000 language-independent core concepts that are expressed in all wordnets. The Domain Ontology consists of a set of topical concepts like “medicine” and “traffic”; unlike the unstructured list of domain labels in the Princeton WordNet, the domain concepts form a hierarchy.

In contrast to the individual wordnets, which are semantic networks with hierarchical relations, the ILI is an unstructured, flat list of lexicalized concepts. Each is represented by a synset and an English definition of the concept expressed by the synset members. The ILI started out as the Princeton WordNet, with each synset being assigned a unique identification number, or “record.” The words and synsets of the languages of EWN were mapped, to the extent possible, onto the synsets in the ILI, and the record identification number was attached to the corresponding word or synset in the target language.

In those cases where a language has one or more words expressing a concept that is not lexicalized in English (i.e., lacking corresponding English words), a new record was created in the ILI with just an identification number but without English lexemes; this record includes a pointer to the synset in the source language. In this way, the ILI came to include, besides WordNet’s synsets, records for all concepts that are lexicalized in one or more EuroWordNet language but not in English. The ILI thus constitutes the superset of all concepts included in all European wordnets.

By means of the records, the ILI mediates among the synsets of the individual languages. Equivalent concepts and words across languages can be determined by referencing the appropriate ILI records.

Maintaining the ILI is a flat list of entries and restricting the encoding of lexical and semantic relations to each of the language-specific wordnets avoids the problem of crosslinguistic mismatches in the patterns of lexicalization and hierarchical organization. For example, Vossen (2004) cites the case of English *container*, which has no counterpart in Dutch. Dutch does have, however, words for specific kinds of containers, like *box*, *bag*, etc. If the ILI had taken over the English WordNet’s hierarchical structure, where *container* is a superordinate of *box*, *bag*, etc., mapping to Dutch would be problematic. Instead, the Dutch wordnet simply maps the Dutch words for *box*, *bag*, etc., to its (Dutch) lexicalized superordinate (implement) and disregards the *container* level.

As the number of wordnets grows, the need for connecting them to one another becomes pressing; moreover, linking should be able not only accurate matching of cross-linguistically encoded concepts but also allow for the expression of meanings that are specific to a language or culture. Fellbaum and Vossen (2007) and Vossen and Fellbaum (2009) propose the creation of a suitable infrastructure dubbed the “Global Grid.”

Interconnected multilingual wordnets carry tremendous potential for crosslinguistic NLP applications and the study of universal and language-specific lexicalization patterns.

10.15 Global WordNets

Currently, wordnet databases exist for several dozen languages (see Singh, 2002; Sojka et al., 2004, 2006) and new ones are being developed.⁴ Virtually all wordnet developments follow the methodology of EuroWordNet described earlier. But wordnets for typologically distinct languages pose novel challenges especially with respect to the notions “concept” and “word,” which must be defined to determine synsets and synset membership. One challenge is the morphology of agglutinative languages like Turkish, Estonian, Tamil and Basque (Bilgin et al., 2004; Kahusk and Vider, 2002; Thiyagarajan 2002; Agirre et al., 2002), where multiple affixes that carry grammatical and lexical meaning are added to a stem to form a single long “word”. For example, does a diminutive formed via an affix express a concept distinct from the base form or are they merely lexical variants? Diminutives are arguably independent words, and they could be included in a wordnet as such, with a pointer expressing a “diminutive” relation to the base form.

Even more challenging are languages like Hebrew and Arabic, where words are generated from a triconsonantal root that constitutes a kind of “super-concept” but that does not have lexical status itself; words whose meanings share the core meaning of the root are derived from it via the addition of vowels (Black et al., 2006).

For Chinese, Wong and Pala (2004) propose to exploit the semantics inherent in the Chinese writing system. A character typically consists of two radicals, one of which carries meaning while the other indicates the pronunciation. Characters and the concepts they express can be grouped and related to one another based on the meaning-carrying radical, at least at the top and middle level of the hierarchies.

10.16 WordNet as a Tool for Natural Language Processing

WordNet’s design and electronic format have proved useful for a wide range of Natural Language Processing (NLP) applications, including mono- and crosslinguistic information retrieval, question-answering systems, and machine translation.

⁴See the website of the Global WordNet Organization, <http://www.globalwordnet.org>

All these tasks face the challenge of word sense identification posed by lexical polysemy. Statistical approaches can identify the context-intended sense in many cases but are limited. WordNet facilitates alternative or complementary symbolic approaches to word sense discrimination, as it allows automatic systems to detect and measure the semantic relatedness of polysemous words that co-occur in a context.⁵

10.17 Conclusions

WordNet represents a new approach – made possible by its electronic format – towards revealing the systematic ways in which a language maps concepts onto words. WordNet deliberately focuses on the lexicon, but its rigid structure and representation of upper-level words and concepts have sometimes invited its comparison to an ontology, a language-independent knowledge structure. Mapping concepts in formal ontologies to synsets in wordnets maintains that distinction and sheds light on concept-word mapping patterns.

Crosslinguistic wordnets show significant overlap at the top levels but diverge on the middle and lower levels, often due to language-specific lexicalization patterns. Further research on WordNet and the development of wordnets in genetically unrelated and typologically diverse languages should advance our understanding of universal and language-specific conceptual and lexical structure.

Acknowledgements Work partially supported by contracts from The Disruptive Technology Office, REFLEX, and the U.S. National Science Foundation.

References

- Agirre E., O. Ansa, X. Arregi, J.M. Arriola, A. Diaz de Ilarraza, E. Pociello, and L. Uria. 2002. Methodological issues in the building of the Basque Word-Net: Quantitative and qualitative analysis. In: Proceedings of the First International WordNet Conference, Mysore, India: CIIL.
- Bilgin, O., O. Cetimo, and K. Oflaver. 2004. Building a word-net for Turkish. *Romanian Journal of Information Science and Technology* 7:1–2, 163–172.
- Black, W., S. Elkateb, H. Rodriguez, M. Alkhalifa, P. Vossen, A. Pease, M. Bertran, and C. Fellbaum. 2006. The Arabic WordNet project. *Proceedings of the Conference on Lexical Resources in the European Community*, Genoa, Italy.
- Baker, C., C.J. Fillmore, and J. B. Lowe. 1998. The Berkeley FrameNet project. *Proceedings of the COLING-ACL Montreal*, Canada.
- Chaffin, R. 1992. The concept of a semantic relation. In *Frames, fields, and contrasts: New essays in semantic and lexical organization*, eds. A. Lehrer and E.F. Kittay, 253–288. Hillsdale, NJ: Lawrence Erlbaum.
- Collins, A. M., and M. R. Quillian. 1969. Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* 8:240–247.
- Cruse, A. 1986. *Lexical semantics*. Cambridge: Cambridge University Press.
- Fellbaum, C. ed. 1998a. *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
- Fellbaum, C. 1998b. A semantic network of english verbs. In ed. C. Fellbaum, 69–104.

⁵For a bibliography on WordNet-based work see <http://lit.csci.unt.edu/wordnet>.

- Fellbaum, C. 2000. Autotroponymy. In *Polysemy*, eds. Y. Ravin, and C. Leacock, 52–67. Cambridge, MA: Cambridge University Press.
- Fellbaum, C., and J. Kegl. 1989. Taxonomic structure and object deletion in the english verbal system. In Proceedings of the 6th Eastern States Conference on Linguistics, eds. K. deJong, and Y. No, 94–103. Columbus, OH: Ohio State University.
- Fellbaum, C., and G. A. Miller. 2003. Morphosemantic links in Word-Net. *Traitement Automatique des Langues* 44(2):69–80.
- Fellbaum, C. and P. Vossen. 2007. Connecting the universal to the specific: Towards the global grid. In Proceedings of the First International Workshop on Intercultural Communication, Kyoto University, Japan. Reprinted in Intercultural Collaboration: First International Workshop. LNCS 4568, eds. T. Ishida, S.R. Fussell, and P.T. J. M. Vossen, 1–16. New York: Springer.
- Gangemi, A., N. Guarino, C. Masolo, A. Oltramari, and L. Schneider. 2002a. *Sweetening ontologies with DOLCE*. In Proceedings of the 13th International Conference on Knowledge Engineering and Knowledge Management (EKAW02), 166, LNCS 2473.
- Gangemi, A., N. Guarino, A. Oltramari, and S. Borgo. 2002b. Cleaning-up WordNet's top-level. In Proceedings of the First International WordNet Conference, Mysore, India: CIIL.
- Gross, D., U. Fischer, and G.A. Miller. 1989. The organization of adjectival meanings. *Journal of Memory and Language* 28:92–106.
- Kahusk, N., and K. Vider. 2002. Estonian WordNet benefits from word sense disambiguation. In Proceedings of the First Global WordNet Conference, 26–31. Mysore, India: Central Institute of Indian Languages. .
- Kilgarriff, A. 2000. Review of WordNet. *Language* 76:706–708.
- Lin, D. 1999. Review of WordNet: An electronic lexical database. *Computational Linguistics* 25:2.
- Miller, G.A. 1995. WordNet: A lexical database for english. *Communications of the ACM* 38: 39–41.
- Miller, G.A. 1998. Nouns in WordNet. In *Word-Net: An electronic lexical database*, ed. C. Fellbaum, Cambridge, MA: MIT Press.
- Miller, G. A., and C. Fellbaum. 1991. *Semantic networks of english*. Cognition, special issue, eds. B. Levin, and S. Pinker, 197–229. Reprinted in *Lexical and Conceptual Semantics*, eds. B. Levin, and S. Pinker, 197–229. Cambridge, MA: Blackwell.
- Miller, G.A., and F. Hristea. 2004. WordNet nouns: Classes and instances. *Computational Linguistics* 32:1.
- Miller G., C. Leacock, T. Randee, R. Bunker. 1993. A semantic concordance. In Proceedings of the 3rd DARPA Workshop on Human Language Technology, Plainsboro, NJ, 303–308.
- Miller, K. 1998. Modifiers in WordNet. In *Word-net: An electronic lexical database*. ed. C. Fellbaum, Cambridge, MA: MIT Press.
- Niles, I., and A. Pease. 2001. Towards a standard upper ontology. In Proceedings of the 2nd International Conference on Formal Ontology in Information Systems, Ogunquit, Maine.
- Niles, I., and A. Pease. 2003. Linking lexicons and ontologies: Mapping WordNet to the suggested upper merged ontology. In Proceedings of the IEEE International Conference on Information and Knowledge Engineering, 412–416.
- Niles, I. and A. Terry. 2004. The MILO: A general-purpose, mid-level ontology. In Proceedings of the International Conference on Information and Knowledge Engineering, Las Vegas, Nevada.
- Oltramari, A., A. Gangemi, N. Guarino, and C. Masolo. 2002. Restructuring WordNet's top-level: The ontoclean approach. In Proceedings of LREC, Las Palmas, Spain.
- Pease, A. 2003. The sigma ontology development environment. Working Notes of the IJCAI-2003 Workshop on Ontology and Distributed Systems. Proceedings of CEUR 71.
- Pease, A., and C. Fellbaum. 2009. Formal ontology as interlingua: The SUMO and WordNet Linking Project. In *Ontologies and the Lexicon*, eds. C.-R. Huang, N. Calzolari, A. Gangemi, A. Lenci, A. Oltramari, and L. Prévot, 29–43. Cambridge Studies in Natural Language Processing. Cambridge: Cambridge University Press.
- Sojka, P., K. Pala, P. Smrz, C. Fellbaum, and P. Vossen, eds. 2004. Proceedings of the Second International WordNet Conference. Brno, Czech Republic: Masaryk University.

- Sojka, P., K.-S. Choi, C. Fellbaum, and P. Vossen. 2006. *Proceedings of the third global WordNet conference*. Brno, Czech Republic: Masaryk University.
- Thiyagarajan, S., S. Arulmozi, S. Rajendran. 2002. Tamil WordNet. In *Proceedings of the First Global WordNet Conference*, Mysore, India: CIIL.
- Vossen, P. 1998. ed. *EuroWordNet*. Dordrecht: Kluwer.
- Vossen, P. 2004. EuroWordNet: A multilingual database of autonomous and language-specific wordnets connected via an interlingual index. *International Journal of Lexicography* 17(2):161–173.
- Vossen, P., and C. Fellbaum. 2009. Universals and idiosyncracies in multilingual wordnets. In *Multilingual lexical resource*, ed. H. Boas, Hans, 319–345. Berlin: de Gruyter.
- Wong, S.H.S., and K. Pala. 2004. Chinese characters and top ontology in EuroWordNet. In *Proceedings of the First Global WordNet Conference*, ed. U.N. Singh, 224–233. Mysore, India: CIIL.