

## THE BOUNDARY BETWEEN WORD KNOWLEDGE AND WORLD KNOWLEDGE

Judy Kegl  
The Program in Linguistics  
Princeton University  
Princeton, New Jersey 08544

I will focus my comments on linguistic considerations concerning the interrelation between words and world representation and will argue that word knowledge must be kept distinct from world knowledge. Knowing everything about a referent object or event with which a word is associated is not enough to allow one to use a word appropriately. World knowledge must be supplemented and constrained by linguistic knowledge to yield an appropriate account of word knowledge.

Let's start with a relatively straightforward example, one which is now so commonplace that I am unsure of who to cite for its introduction into the literature. Consider the word *dance/danser* in both English and French. These cognates might at first glance appear to be almost interchangeable glosses for one another. Certainly, when looking at the event we think of as dancing, both English and French speakers see and experience the same physical event. Details of the exact instantiation of dancing might vary from culture to culture, but I think we would agree that we all share the same concept of what constitutes a dancing event. Yet speakers of English and French do not use their "words" for this event in parallel ways. In English, for example, we use the verb *dance* to refer to both a translatory and a nontranslatory event (where translatory means to move along a path from one place to another). So, for example, the English sentence *John and Mary danced across the room* is ambiguous. It can mean either 'John and Mary went from one end of the room to the other while dancing' or 'John and Mary were located at the other end of the room dancing.' French, on the other hand, allows only the nontranslatory reading where John and Mary are at some location engaged in the act of dancing. French speakers have no problem describing the translatory event of dancing from one end of the room to the other, they just don't use the word *danse* in this context. They speak of crossing the room while dancing. Somewhere along the road between French and English world knowledge and word knowledge seem to have diverged. What's interesting about such examples is that this divergence isn't a freak occurrence. The divergence seen with *dance* and *danse* extends to a whole class of related verbs. In other words, the behavior of these verbs is a reflex of some language-specific regularities.

However, even though there exists a considerable amount of freedom with regard to what aspects of our world knowledge of an object or event are encoded in its linguistic alter ego (the word), the range of possibilities is still severely constrained. Not everything we know about an object or an event is a candidate for inclusion in its linguistic representation. For example, color is a salient property we associate with the objects around us, but no language in the world exhibits a linguistic classifier system based on color. Shape, solidity and flexibility, on the other hand, are at the heart of a large majority of the world's classifier systems. Time and space seem to have equal status in the world of conceptual knowledge, yet languages prefer to speak of time in spatial terms (at 6:00, from Tuesday to Thursday, toward evening, in 1976, Winter is fast ap-

proaching) and not *vice versa*. From the time of the Byzantine Planudes through the writings of Hjelmslev (1935) and even into the more recent work of Gruber (1976) and Jackendoff (1983), the claim has been made that motion and location play a special role in the organization of language. Even abstract verb classes (verbs of cognition, perception, emotion, *etc.*) involve figurative extensions of basic motion/location relations (to feel down, to think something through). When we stop to think about it, many possibilities for classification come to mind, but the world's languages seem to select their options from a very restricted set. Determining these preferred encoding strategies across languages is one way to get at a linguistically relevant set of semantic primitives. The relevant linguistic phenomena to examine are verb classes (cognition, perception, emotion, motion/location, change of state, bodily care, *etc.*), thematic roles (source, goal, theme) and predicate argument structure, mappings between semantic roles (agent, patient) and grammatical relations (subject, object), and grammatical functions such as causation. All of these phenomena straddle the boundary between linguistic conceptual structure and conceptual structure in general.

Acquisition work by the Clarks (E. Clark 1973, H. Clark 1973) presents some ties between perception and language and raises some interesting issues concerning the interface between perceptual and linguistic knowledge, and the mechanism by which a child might use perceptual knowledge as a bootstrapping device to link words and world knowledge into an already existent innate linguistic knowledge.

Moving to evidence from language typology, we find that even though grammatical relations and semantic roles (agent, patient) are evidently part of universal grammar, the link between them is subject to variation and must be determined on the basis of language-specific evidence. Otherwise, how could Nominative/Accusative languages (languages where the object is the marked case) map agent to subject and patient to object, whereas Ergative languages (those where the subject of a transitive is the marked case) do the reverse, mapping agent to object and patient to subject. This is a simplified statement of the facts (see Levin 1983 for more detail), but it is explicit enough to make my point here, which is that certain links between world knowledge and linguistic knowledge are in a very constrained sense arbitrary.

Recent work on lexical conceptual structures coming out of the Lexicon Project at MIT casts some of the words and world representation issues in a new light. Hale and his colleagues (Hale and Laughren 1983, Hale and Keyser 1986) have proposed a level of representation called lexical conceptual structure (LCS) which follows from and extends Jackendoff's work on predicate decomposition. This work involves breaking the definition of a predicator into a series of subcomponents of meaning, subcomponents which prove to be relevant to the predicate argument structure of these lexical items. The LCS offers insight into the mapping between lexical representations and syntactic configurations. In my opinion, when we speak of lexical semantic representation we must keep in mind that the semantic representation of a word relevant to its functioning as a linguistic entity is separate from the much wider range of information which must be considered to be part of its meaning. For example, connotation is relevant to the meaning of a word but it is not part of its lexical semantic representation.

Consider another crosslinguistic discussion. The meaning of a word like *thirst* has the potential for yielding multiple LCSs only some of which might be relevant to its linguistic behavior.

Take the sensation of being thirsty, which we share with all other human beings, and let's agree for the sake of argument that we can all share the same concept of a thirst event. When we look at how that event is spoken of across languages, we find that it is linguistically encoded in a variety of different ways. In some languages the lexical semantic representation for 'to thirst' might be nontranslatory (I have thirst, Thirst is at me, I am at thirst (in a state of being thirsty)); or it might be translatory (I go to thirst, I am getting to be thirsty, Thirst has come upon me, *etc.*). Notice that even within the set of translatory *versus* nontranslatory options the thematic roles assigned to *thirst* can vary. In some instances it serves as the theme (the entity which moves or is located), and in other instances it serves as an anchorpoint which the experiencer moves to (goal) or is located at (location). I am sure if we looked hard enough we would also find examples where a source is involved (I am running from thirst, Thirst comes from me). All of the above figure/ground relations and translatory/nontranslatory options are compatible with the meaning of *thirst*, and they all pick from the same set of building blocks comprised of motion/location relations and thematic roles. Still, which encoding to use as the basis of a LCS for a verb of thirsting is a language-specific choice. And, certain specific syntactic and semantic consequences will follow from that choice--similar in nature to the syntactic and semantic differences between *dance* and *danseur*.

Actually, the work on LCSs raises two even more basic questions: What is a word? And, what does it mean to know a word? A word is not the header associated with a lexical entry. That is simply a label associated with a given word. A word is a more abstract entity which is probably best thought of as a LCS plus information concerning what components of that LCS are projected into the syntax, its predicate argument structure. The label *move* in English, for example, is associated with at least two words, one nontranslatory (I saw it move), and the other translatory (She moved to the other side of the room). In fact, some languages give distinct labels to these two words.

Knowing a word entails not only knowing what it means, but also knowing how to use it. Part of knowing a verb is knowing the various ways it can realize its arguments, and, if it permits more than one way, what sets each of the alternatives apart. Examination of this aspect of lexical knowledge gets us into the nitty gritty of what word representation is all about. It also brings to light some of the more perplexing problems which must be faced in the process of developing "reverse engineering" techniques for extracting information from dictionaries. I will mention not only the problems, but also the potential benefits to be gained from systematic examination of learner's dictionaries and large corpora of texts. Atkins, Kegl and Levin (in press) discusses the explicit and implicit information to be found in learner's dictionaries such as the *Longman Dictionary of Contemporary English* (LDOCE), the *Oxford Advanced Learner's Dictionary of Contemporary English* (OALD) and the *Collins English-Learner's Dictionary*. In this paper, we pointed out complex form/meaning interdependencies which can only be extracted from existing dictionaries by detailed examination of converging information provided in the three central components of a dictionary entry: the syntactic code, the sense information and the example sentences.

Consider the problem. Many verbs exhibit transitivity alternations where they appear as either transitive or intransitive. See the examples below which are discussed in the Atkins, Kegl and Levin paper:

Indefinite (Unspecified) Object Alternation:

- (1) a. Mike ate the cake.
- b. Mike ate. (=Mike ate food or a meal.)

Characteristic Property Alternation:

- (2) a. That dog bites people.
- b. That dog bites. (=That dog is a biter.)

Reciprocal Alternation:

- (3) a. Anne met Cathy.
- b. Anne and Cathy met. (=Anne and Cathy met each other.)

Reflexive Alternation:

- (4) a. Jill dressed herself hurriedly.
- b. Jill dressed hurriedly. (=Jill dressed herself hurriedly.)

Causative/Inchoative (also Anti-Causative or Ergative) Alternation:

- (5) a. Janet broke the cup. (=Janet caused the cup to break.)
- b. The cup broke. (how the cup broke is left unspecified)

Instrumental Alternation:

- (6) a. The scissors won't cut the denim.
- b. The scissors won't cut.

The preceding examples not only participate in transitivity alternations, but the interpretation assigned to their intransitive forms and the thematic role assigned to their subjects varies according to the verb class to which each belongs. When examining pairs such as those above, it quickly becomes clear that a simple transitive/intransitive distinction will not suffice. There is a complex interaction between verb class, sense and transitivity which must be recognized. Work on extracting significant lexical information from machine-readable dictionaries such as the research described by Boguraev (1986) requires extremely complex computations on a data base not originally designed with such a use in mind. Although learner's dictionaries are phenomenally comprehensive, they are not as consistent in their presentation of the lexical data as one might hope. Therefore, a lot of cleaning up needs to be done before even the best machine-readable dictionary can reveal the form/meaning interdependencies mentioned above. The project is, however, well worth the effort. One suggestion I would like to make is that we not confine ourselves to decoding simply the syntactic code and sense portions of the dictionary. The example sentences have a lot to offer. Contrary to their role in etymological dictionaries or more general purpose dictionaries where they serve to document the chronological occurrences of a lexical item or its attested occurrences in literary texts, example sentences in learner's dictionaries are another form of code. The sentences which exemplify certain verb classes or types of transitivity alternations are parallel in form, so much so that one can often determine the possibilities for syntactic patterns and verb class membership from the examples alone. And, frequently the examples cue the user into usage patterns that the entry itself has failed to include in its metalinguistic components. Finally, a pair of sentences and their associated interpretations are frequently the most effective means of conveying cross-classificatory information to a user. Even if a dictionary were able to exhaustively convey this information in its code, it is not clear that such information

would be easily understood by the user. Furthermore, the more we develop tools for extracting dictionary information from carefully chosen example sentences, the closer we will get to extracting such information from unrestricted texts.

Determining verb class membership and the syntactic and semantic properties of a given lexical item also depends upon world knowledge that the user brings to the dictionary, knowledge which is pre-assumed by the sense definitions. In analyzing dictionary entries in LDOCE and OALD, we found many inconsistencies in the presentation of lexical entries which only became obvious when we examined what one would get from the explicit and implicit regularities presented alone, without certain common sense and experiential knowledge we are assumed to bring to the task. Unless the "reverse engineering device" also comes to the task with such assumed knowledge, we won't reap the maximum benefits from this dictionary research.

A final postscript to this discussion concerns current efforts to develop a theory neutral lexicon which can be used by a variety of parsing and generation programs. Such a goal can be achieved, but I think it is important to keep in mind how that end product will be achieved and what it has to offer theory-based parsers and generation systems. The neutral lexicon must be reductionist in nature. It can only be arrived at by continually breaking into subcomponents those units currently utilized by well-articulated linguistic theories. Then we will still need a series of overlay systems which will combine those commonly shared units into the chunks made reference to in each individual framework. Thus it seems that building a neutral lexicon requires a well articulated lexicon from each theory purporting to use this neutral lexicon and a well articulated set of rules to translate the neutral lexicon into a form the framework can utilize. On the optimistic side, although we work hard to maintain our individual identity, most linguists do seem to agree on most of the basic units of analysis as is evidenced by the fact that we continually translate the data from one theory into another in order to argue against it or in order to use the basic facts in support of our own analyses. Work toward a neutral lexicon will, if nothing else, lead us to a better understanding of the subsets of data successfully handled in different frameworks and to a more precise characterization of those points on which we agree and disagree.

## REFERENCES

- Atkins, B.T., J. Kegl and B. Levin (in press) "Explicit and Implicit Information in Dictionaries." In *Proceedings of the Conference on Advances in Lexicology*, Center for the New Oxford English Dictionary, Waterloo.
- Boguraev, B. (1986) "Machine-Readable Dictionaries and Research in Computational Linguistics." In Walker, *et al.*, eds., *Proceedings of the Workshop on Automating the Lexicon*, Grosseto, Italy.
- Carver, D.J. (1974) *et al.*, eds, *Collins English Learner's Dictionary*, Collins, Ltd., London.
- Clark, E. (1973) "Non-Linguistic Strategies and the Acquisition of Word Meaning," *Cognition*, vol. 2, pp. 161-182.
- Clark, H. (1973) "Space, Time, Semantics and the Child." In T.E. Moore, *Cognitive Develop-*

- ment and the Acquisition of Language*, Academic Press, New York.
- Gruber, J. (1976) *Lexical Structures in Syntax and Semantics*, North-Holland, Amsterdam.
- Hale, K. and S.J. Keyser (1986) "A View from the Middle," *Lexicon Working Papers 10*, Center for Cognitive Science, MIT, Cambridge, MA.
- Hale, K. and M. Laughren (1983) "Preface to Dictionary Entries of Verbs." Unpublished manuscript, MIT, Cambridge, MA.
- Hjelmslev, L. (1935) "La Categorie des Cas," *Acta Jutlandica*, vol. 7, pp. i-xiii, 1-184.
- Hornby, A.S. and A.P. Cowie, eds. (1974, first edition 1948) *Oxford Advanced Learner's Dictionary of Current English*, Oxford University Press, Oxford.
- Jackendoff, R. (1983) *Semantics and Cognition*, MIT Press, Cambridge, MA.
- Levin, B. (1983) *On the Nature of Ergativity*. Unpublished doctoral dissertation, MIT, Cambridge, MA.
- Proctor, P., ed. (1978) *Longman Dictionary of Contemporary English*, Longman, London.