Language Games,
A Foundation for Semantics and Ontology

John F. Sowa

The issues raised by Wittgenstein's language games are fundamental to any theory of semantics, formal or informal. Montague's view of natural language as a version of formal logic is at best an approximation to a single language game or a family of closely related games. But it is not unusual for a short phrase or sentence to introduce, comment on, or combine aspects of multiple language games. The option of dynamically switching from one game to another enables natural languages to adapt to any possible subject from any perspective for any humanly conceivable purpose. But the option of staying within one precisely defined game enables natural languages to attain the kind of precision that is achieved in a mathematical formalism. To support the flexibility of natural languages and the precision of formal languages within a common framework, this article drops the assumption of a fixed logic. Instead, it proposes a dynamic framework of logics and ontologies that can accommodate the shifting points of view and methods of argumentation and negotiation that are common during discourse. Such a system is necessary to characterize the open-ended variety of language use in different applications at different stages of life — everything from an infant learning a first language to the most sophisticated adult language in science and engineering.

This is a preprint of an article that appeared as Chapter 2 in *Game Theory and Linguistic Meaning,* edited by Ahti-Veikko Pietarinen, Elsevier, 2007, pp. 17-37.

1. The Infinite Flexibility of Natural Languages

Natural languages are easy to learn by infants, they can express any thought that any adult might ever conceive, and they are adapted to the limitations of human breathing rates and short-term memory. The first property implies a finite vocabulary, the second implies infinite extensibility, and the third implies a small upper bound on the length of phrases. Together, they imply that most words in a natural language will have an open-ended number of senses — ambiguity is inevitable. Charles Sanders Peirce and Ludwig Wittgenstein are two philosophers who understood that vagueness and ambiguity are not defects in language, but essential properties that enable it to express anything and everything that people need to say. This article takes these insights as inspiration for a system of metalevel reasoning, which relates the variable meanings of a finite set of words to a potentially infinite set of concept and relation types, which are used and reused in dynamically evolving lattices of theories, which may be expressed in an open-ended variety of logics.

At the beginning of his career, Wittgenstein, like many of the early researchers in artificial intelligence, thought he had found the key to solving the problems of understanding language and reasoning. In his first book, the *Tractatus Logico-Philosophicus*, he presented an elegant view of semantics that directly or indirectly inspired the theories of formal semantics and knowledge representation that were developed in the 20th century: an elementary proposition expresses an atomic fact about a state of affairs (Sachverhalt), which consists of a configuration of objects (Verbindung von Gegenständen); a compound proposition is a Boolean combination of elementary propositions; everything in the world can be described by some proposition, elementary or compound; and everything that can be said can be clearly expressed by some proposition about such configurations. His conclusion was the famous one-sentence Chapter 7, which conveniently dismissed all exceptions: "Whereof one cannot speak, thereof one must be silent."

The *Tractatus* inspired Rudolf Carnap's version of logical positivism and Alfred Tarski's model-theoretic semantics. One of Tarski's students, Richard Montague, extended model theory to intensional verbs, such as *believe*, *want*, or *seek*. Montague's grammar (1970) mapped a *fragment* of English to models with an elaborate construction of multiple worlds instead of Wittgenstein's single world. Around the same time, Woods (1968, 1972) and Winograd (1972) implemented model-theoretic systems for talking about moon rocks and the blocks world. Winograd's thesis adviser, Marvin Minsky, was also a technical adviser for the movie *2001, A Space Odyssey,* which featured the HAL 9000, a computer that not only spoke and understood English, but could also read lips, interpret human intentions, and conceive plans to thwart them. When the movie appeared in 1968, Minsky claimed it was a conservative prediction about AI technology in 2001.

Although Wittgenstein and Winograd had a strong influence on later developments, both of them became disillusioned about a decade after their early successes. After Wittgenstein published his first book, which he believed had solved all the solvable problems of philosophy, he went to teach school in an Austrian mountain village. Unfortunately, his pupils didn't think or speak the way his theory predicted. It was impossible to find any truly atomic facts that could not be further analyzed or viewed from an open-ended number of different perspectives. Winograd also became discouraged by the difficulty of generalizing and extending his early system, and he later published a harsh critique of his own and other methods for translating natural language to logic (Winograd & Flores 1986). Today, no AI system has any ability that can remotely compare to the HAL 9000, and textbooks based on Montague's approach are illustrated with toy examples that more closely resemble Montague's fragment than the English that anybody actually reads, writes, or speaks.

The precision of logic is valuable, but what logic expresses so precisely may have no relationship to what was intended or required. A formal specification that satisfies the person who wrote it might not satisfy the users' requirements. Engineers summarize the problem in a pithy slogan: "Customers never know what they want until they see what they get." More generally, the precision and clarity that are so admirable in the final specification of a successful design are the result of a lengthy process of trial, error, and revision. In most cases, the process of revision never ends until the system is obsolete.

Unlike formal languages, which can only express the finished result of a lengthy analysis, natural languages can express every step from an initially vague idea to the final specification. During his career as an experimental physicist and a practicing engineer, Peirce learned the difficulty of stating any general principle with absolute precision:

> It is easy to speak with precision upon a general theme. Only, one must commonly surrender all ambition to be certain. It is equally easy to be certain. One has only to be sufficiently vague. It is not so difficult to be pretty precise and fairly certain at once about a very narrow subject. (CP 4.237)

This quotation summarizes the futility of any attempt to develop a precisely defined ontology of everything, but it offers two useful alternatives: an informal classification, such as a thesaurus or terminology, and an open-ended collection of formal theories about narrowly delimited subjects. It also raises the questions of how and whether these resources might be used as a bridge between informal natural language and formally defined logics and programming languages.

Even if an ideal semantic representation were found, it would not answer the question of how any system, human or machine, could learn and use the representation. Children rapidly learn to associate words with the things and actions they see and do without analyzing them into atomic facts or evaluating Montague's functions from possible worlds to truth values. As an example, the following sentence was spoken by a child named Laura at age 34 months (Limber 1973):

> *When I was a little girl, I could go "geek, geek" like that;*
> *but now I can go "This is a chair."*

In this short passage, Laura combined subordinate and coordinate clauses, past tense contrasted with present, the modal auxiliaries *can* and *could*, the quotations "geek, geek" and "This is a chair," metalanguage about her own linguistic abilities, and parallel stylistic structure. The difficulty of simulating such ability led Alan Perlis to remark "A year spent in artificial intelligence is enough to make one believe in God" (1982).

2. Wittgenstein's Alternative

Although Wittgenstein criticized his earlier theory of semantics and related theories by Frege and Russell, he did not reject everything in the *Tractatus*. He continued to have a high regard for logic and mathematics, and he taught a course on the foundations of mathematics, which turned into a debate between himself and Alan Turing. He also retained the picture theory of the Tractatus, which considered the relationships among words in a sentence as a *picture* (Bild) of relationships in the world. What he abandoned, however, was the claim that there exists a unique decomposition of the world into atomic facts and a privileged vantage point for taking pictures of those facts. A chair, for example, is a simple object for someone who wants to sit down; but for a cabinet maker, it has many parts that must be carefully fit together. For a chemist developing a new paint or glue, even the wood is a complex mixture of chemical compounds, and those compounds are made up of *atoms*, which are not really atomic after all. Every one of those views is a valid picture of a chair for some purpose.

In the *Philosophical Investigations*, Wittgenstein showed that ordinary words like *game* have few, if any common properties that characterize all their uses. Competition is present in ball games, but absent in solitaire or ring around the rosy. Organized sport follows strict rules, but not spontaneous play. And serious games of life or war lack the aspects of leisure and enjoyment. Instead of unique defining properties, games share a sort of *family resemblance*: baseball and chess are games because they resemble the family of activities that people call games. Except for technical terms in mathematics, Wittgenstein maintained that most words are defined by family resemblances. Even in mathematics, the meaning of a symbol is its use, as specified by a set of rules or axioms. A word or other symbol is like a chess piece, which is not defined by its shape or physical composition, but by the rules for using the piece in the game of chess. As he said,

> There are *countless* — countless different kinds of use of what we call 'symbols,' 'words,' 'sentences.' And this multiplicity is not something fixed, given once and for all; but new types of language, new language games, as we may say, come into existence, and others become obsolete and get forgotten. (§23)

As examples of language games, he cited activities in which the linguistic component is unintelligible outside a framework in which the nonlinguistic components are the focus. A child or a nonnative speaker who understood the purpose of the following games could be an active participant in most of them with just a rudimentary understanding of the syntax and vocabulary:

> Giving orders, and obeying them; describing the appearance of an object, or giving its measurements; constructing an object from a description (a drawing); reporting an event; speculating about an event; forming and testing a hypothesis; presenting the results of an experiment in tables and diagrams; making up a story, and reading it; play acting; singing catches; guessing riddles; making a joke, telling it; solving a problem in practical arithmetic; translating from one language into another; asking, thanking, cursing, greeting,

> praying. (§23)

Only the game of describing an object could be explained in the framework of the *Tractatus*. Wittgenstein admitted that it could not explain its own language game: "My propositions are elucidatory in this way: he who understands me finally recognizes them as senseless..." (6.54). The theory of language games, however, is capable of explaining the language game of writing a book about anything, including language games.

In his later work, Wittgenstein faced the full complexity of language as it is used in science and everyday life. Instead of the fixed boundaries defined by necessary and sufficient conditions, he used the term *family resemblances* for the "complicated network of overlapping and criss-crossing similarities" (1953, §66) in which vagueness is not a defect:

> One might say that the concept 'game' is a concept with blurred edges. — "But is a blurred concept a concept at all?" — Is an indistinct photograph a picture of a person at all? Is it even always an advantage to replace an indistinct picture with a sharp one? Isn't the indistinct one often exactly what we need?
>
> Frege compares a concept to an area and says that an area with vague boundaries cannot be called an area at all. This presumably means that we cannot do anything with it. — But is it senseless to say: "Stand roughly (ungefähr) there"? (§71).

Frege's view is incompatible with natural languages and with every branch of empirical science and engineering. With their background in engineering, Peirce and Wittgenstein recognized that all measurements have a margin of error or granularity, which must be taken into account at every step from design to implementation. The option of vagueness enables language to accommodate the inevitable vagueness in observations and the plans that are based on them.

After a detailed analysis of the *Tractatus*, Simons (1992) admitted that Wittgenstein's later criticisms are valid: "We might say that not everything *we* say *can* be said clearly" (p. 357). But he was not ready to adopt language games as the solution: Wittgenstein "became a confirmed — some, including myself, would say *too* confirmed — believer in the messiness of things." Yet things really are messy. As Eugene Wigner (1960) observed, "the unreasonable effectiveness" of mathematics for representing the fundamental principles of physics is truly surprising. The basic equations, such as F=*ma*, are deceptively simple; even their relativistic or quantum mechanical extensions can be written on one line. The messiness results from the application of the simple equations to the enormous number of atoms and molecules in just a tiny speck of matter. When applied to the simplest living things, such as a bacterium, even the fastest supercomputers are incapable of solving the equations. In any practical calculation, such as predicting the weather, designing a bridge, or determining the effects of a drug, drastic approximations are necessary. Those approximations are always tailored to domain-dependent special cases, each of which resembles a mathematical variant of what Wittgenstein called a language game. In fact, he said "We can get a rough picture of [the language games] from the changes in mathematics" (§23).

Although Wittgenstein's ideas are highly suggestive, his definitions are not sufficiently precise to enable logicians to formalize them. Some confusion is caused by the English term *language game*, which suggests a kind of competition that is not as obvious in the original German *Sprachspiel*. Perhaps a better translation might be *language play* or, as Wittgenstein said, the language used with a specific type of activity in a specific form of life (Lebensform). Hattiangadi (1987) suggested that the meaning of a word is the set of all possible *theories* in which it may be used; each theory would characterize one type of activity and the semantics of the accompanying language game. The term *sublanguage*, which linguists define as a semantically restricted dialect (Kittredge & Lehrberger 1982), may be applied to a family of closely related language games and the theories that determine their semantics. The crucial problem is to determine how the members of such families are related to one another, to the members of other families, and to the growing and changing activities of the people — children and adults — who learn them, use them, and modify them.

3. Models of Language

Any theory of language should be simple enough to explain how infants can learn language and powerful enough to support sophisticated discourse in the most advanced fields of science, business, and the arts. Some formal theories have the power, and some statistical theories have the simplicity. But an adequate theory must explain both and show how a child can grow from a simple stage to a more sophisticated stage without relearning everything from scratch: each stage from infancy to adulthood adds new skills by extending, refining, and building on the earlier representations and operations.

During the second half of the 20th century, various models of language understanding were proposed and implemented in computer programs. All of them have been useful for processing some aspects of language, but none of them have been adequate for all aspects of language or even for full coverage of just a single aspect:

- **Statistical.** In the 1950s, Shannon's information theory and other statistical methods were popular in both linguistics and psychology, but the speed and storage capacity of the early computers were not adequate to process the volumes of data required. By the end of the century, the vastly increased computer power made them competitive with other methods for many purposes. Their strength is in pattern-discovery methods, but their weakness is in the lack of a semantic interpretation that can be mapped to the real world or to other computational methods.
- **Syntactic.** Chomsky's transformational grammar and related methods dominated linguistic studies in the second half of the 20th century, they stimulated a great deal of theoretical and computational research, and the resulting syntactic structures can be adapted to other paradigms, including those that compete with Chomsky and his colleagues. But today, Chomsky's contention that syntax is best studied independently of semantics is at best unproven and at worst a distraction from a more integrated approach to language.
- **Logical.** By the 1970s, the philosophical studies from Carnap and Tarski to Kripke and Montague led to formal logics with better semantic foundations and reasoning methods than any competing approach. Unfortunately, those methods can only interpret sentences that have been deliberately written in a notation that looks like a natural language, but is actually a syntactic variant of the underlying logic. None of them can generate logical formulas from the language that people speak or write for the purpose of communicating with other people.
- **Lexical.** Instead of forcing language into the mold of formal logic, lexical semanticists study all features of syntax, vocabulary, and context that can cause sentences to differ in meaning. The strength of lexical semantics is a greater descriptive adequacy and a sensitivity to more aspects of meaning than other methods. Its weakness is a lack of an agreed definition of the meaning of 'meaning' that can be related to the world and to computer systems.
- **Neural.** Many people believe that neuroscience may someday contribute to better theories of how people generate and interpret language. That may be true, but the little that is currently known about how the brain works can hardly contribute anything to linguistic theory. Systems called *neural networks* are statistical methods that have the same strengths and weaknesses as other statistical methods, but they have little resemblance to the way actual neurons work.

Each of these approaches is based on a particular technology: mathematical statistics, grammar rules, dictionary formats, or networks of neurons. Each of them ignores those aspects of language for which the technology is ill adapted. For people, however, language is seamlessly integrated with every aspect of life, and they don't stumble over boundaries between different technologies. Wittgenstein's language games do not compartmentalize language by the kinds of technology that produce it, but by subject matter and mode of use. That approach seems more natural, but it raises the question of how a computer could recognize which game is being played, especially when aspects of multiple games are combined in the same paragraph or even the same sentence.

The greatest strength of natural language is its flexibility and power to express any sublanguage ranging from cooking recipes to stock-market reports and mathematical formulas. A flexible syntactic theory, which is also psychologically realistic, is *Radical Construction Grammar* (RCG) by Croft (2001). Unlike theories that draw a sharp boundary between grammatical and ungrammatical sentences, RCG can accept any kind of construction that speakers of a language actually use, including different choices of constructions for different sublanguages:

> Constructions, not categories or relations, are the basic, primitive units of syntactic representation.... the grammatical knowledge of a speaker is knowledge of constructions (as form-meaning pairings), words (also as form-meaning pairings), and the mappings between words and the constructions they fit in. (p. 46)

RCG makes it easy to borrow a word from another language, such as *connoisseur* from French or H_2SO_4 from chemistry, or to borrow an entire construction, such as *sine qua non* from Latin or $x^2+y^2=z^2$ from algebra. In the sublanguage of chemistry, the same meaning that is paired with H_2SO_4 can be paired with *sulfuric acid*, and the constructions of chemical notation can be freely intermixed with the more common constructions of English syntax.

A novel version of lexical semantics, influenced by Wittgenstein's language games and related developments in cognitive science, is the theory of *dynamic construal of meaning* (DCM) proposed by Cruse (2000) and developed further by Croft and Cruse (2004). The fundamental assumption of DCM is that the most stable aspect of a word is its spoken or written sign; its meaning is unstable and dynamically evolving as it is construed in each context in which it is used. Cruse coined the term *microsense* for each subtle variation in meaning as a word is used in different language games. That is an independent rediscovery of Peirce's view: the spelling or shape of a sign tends to be stable, but each interpretation of a sign token depends on its context in a pattern of other signs, the physical environment, and the interpreter's memory of previous patterns. Croft and Cruse showed how the DCM view of semantics could be integrated with a version of RCG, but a more detailed specification is required for a computer implementation.

In surveying the difficulties of language translation, Steiner (1975) observed that the most amazing fact about languages is the multiplicity of radically different means for expressing idiosyncratic views of the world:

> No two historical epochs, no two social classes, no two localities use words and syntax to signify exactly the same things, to send identical signals of valuation and inference. Neither do two human beings. Each living person draws, deliberately or in immediate habit, on two sources of linguistic supply: the current vulgate corresponding to his level of literacy, and a private thesaurus. The latter is inextricably a part of his subconscious, of his memories, so far as they may be verbalized, and of the singular, irreducibly specific ensemble of his somatic and psychological identity. Part of the answer as to whether there can be 'private language' is that aspects of every language act are unique and individual. They form what

> linguists call an ‘idiolect’. Each communicatory gesture has a private residue. The ‘personal lexicon’ in every one of us inevitably qualifies the definitions, connotations, semantic moves current in public discourse. The concept of a normal or standard idiom is a statistically-based fiction (though it may, as we shall see, have real existence in machine translation). The language of a community, however uniform its social contour, is an inexhaustibly multiple aggregate of speech-atoms, of finally irreducible personal meanings.... Thus a human being performs an act of translation, in the full sense of the word, when receiving a speech-message from any other human being. (pp. 47-48)

The multiplicity of unique language forms, which makes translation difficult even for the best human translators, is an even greater challenge for machine translation. Steiner’s remark about a “private thesaurus” for each person’s idiolect and a “statistically-based fiction” for MT is intriguing. It suggests the possibility of supporting artificial idiolects by compiling a thesaurus classified according to the language games the machine is designed to play.

4. Semantic Representations

The hypothesis of a prelinguistic semantic representation is as old as Aristotle:

> Spoken words are symbols of experiences (*pathêmata*) in the psyche; written words are symbols of the spoken. As writing, so is speech not the same for all peoples. But the experiences themselves, of which these words are primarily signs, are the same for everyone, and so are the objects of which those experiences are likenesses. (*On Interpretation* 16a4)

Whether that representation is called experience in the psyche, conceptual structure, language of thought, or natural logic is less important than its expressive power, its topological structure, and the kinds of operations that can be performed with it and on it.

Some representations are designed to support Steiner’s informal “aggregates of speech atoms” or “irreducible personal meanings”, but others force language into a rigid, logic-based framework. From his work as a lexicographer, Peirce realized that symbols have different meanings for different people or for the same person on different occasions:

> For every symbol is a living thing, in a very strict sense that is no mere figure of speech. The body of the symbol changes slowly, but the meaning inevitably grows, incorporates new elements and throws off old ones. (CP 2.222).

But as a mathematician and logician, he also recognized the importance of discipline and fixed definitions: “reasoning is essentially thought that is under self-control” (CP 1.606). Yet self-control is always exercised for a specific purpose. As the purpose changes, the language game changes, and the symbols acquire new meanings.

Although there is no direct way of observing the internal representations, many of their properties can be inferred from the features of natural languages and the kinds of reasoning people express in languages, both natural and artificial. Any adequate theory must directly or indirectly support the following features:

1. Every natural language has a discrete set of meaningful units (words or morphemes), which are combined in systematic ways to form longer phrases and sentences.

2. The basic constructions for combining those units express relational patterns with two or three arguments (e.g., a subject, an optional direct object, and an optional indirect object). Additional

arguments are usually marked by prepositions or postpositions.

3. The logical operators of conjunction, negation, and existence are universally present in all languages. Other operators (e.g., disjunction, implication, and universal quantification) are more problematical.
4. Proper names, simple pronouns, and indexicals that point to something in the text or the environment are universal, but some languages have more complex systems of anaphora than others.
5. Metalanguage occurs in every natural language, and it appears even in ’s speech at age three. It supports the introduction of new words, new syntax, and the mapping from new features to older features and to extralinguistic referents.
6. Simple metalanguage requires at least one level of nested structure. Most major languages support multiple levels of nested clauses and phrases, any of which could contain metalevel comments.

Points #1 and #2 indicate that the semantic representation must support graph-like structures (of which strings and trees are special cases). With the addition of points #3 and #4, it supports a subset of first-order logic. Full FOL would require a flexible syntax that can support nested or embedded constructions, which English and other major languages provide. Points #5 and #6, combined with a flexible syntax, can support highly expressive logical constructions.

As this summary shows, natural languages can express complex logic, but it does not imply that complex logic is a prerequisite for language. Infants successfully use language to satisfy their needs as soon as they begin to utter single words and short phrases. Preschool children learn and use complex language long before they learn any kind of mathematics or formal logic. Although all known natural languages have complex syntax, some rare languages, such as Pirahã (Everett 2005), seem to lack the levels of nesting needed to express full FOL. Everett noted that the Pirahã people have no word for *all* or *every* or even a logically equivalent paraphrase. That limitation would make it hard for them to invent mathematics and formal logic. In fact, their ability to count is limited to the range *one, two, many*.

An adequate semantic representation must be able to cover the full range of language used by people in every culture at every stage of life. In modern science, educated adults create and talk about abstruse systems of logic and mathematics. But the Pirahã show that entire societies can live successfully with at best a rudimentary logic and mathematics. As Peirce observed, logical reasoning is a disciplined method of thought, not a prerequisite for thought — or the language that expresses it.

5. A Wittgensteinian Approach to Language

A semantic approach inspired by Wittgenstein’s language games was developed by Margaret Masterman, one of six students in his course of 1933-34 whose notes were compiled as *The Blue Book* (Wittgenstein 1958). In the late 1950s, Masterman founded the Cambridge Language Research Unit (CLRU) as a discussion group, which became one of the pioneering centers of research in computational linguistics. Her collected papers (Masterman 2005) present a computable version with similarities to Cruse’s DCM:

- A focus on semantics, not syntax, as the foundation for language: “I want to pick up the relevant basic-situation-referring habits of a language in preference to its grammar” (p. 200).
- A context-dependent classification scheme with three kinds of structures: a thesaurus with multiple groups of words organized by areas of use, a *fan* radiating from each word type to each

area of the thesaurus in which it occurs, and dynamically generated combinations of fans for interpreting the word tokens of a text.

- Emphasis on images as a language-independent foundation for meaning with a small number (about 50 to 100) of combining elements represented by ideographs or monosyllables, such as IN, UP, MUCH, THING, STUFF, MAN, BEAST, PLANT, DO.
- Recognition that analogy and metaphor are fundamental to the creation of novel uses of language, especially in the most advanced areas of science. In electromagnetism, for example, Maxwell's elegant mathematics is the culmination of a lengthy process that began with Faraday's vague analogies about lines of force.

Figure 1 shows a fan for the word *bank* with links to each area of Roget's *Thesaurus* in which the word occurs (p. 288). The numbers and labels identify areas in the thesaurus, which, Masterman claimed, correspond to "Neo-Wittgensteinian families".

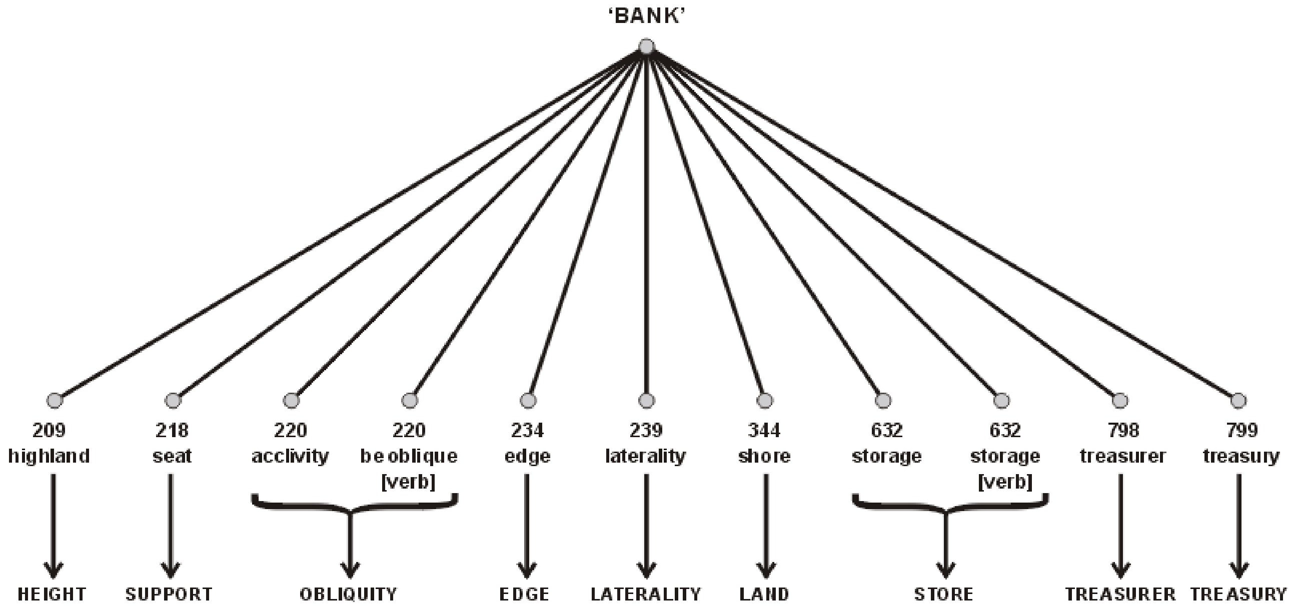

Figure 1: A word fan for *bank*

To illustrate the use of word fans, Masterman analyzed the phrases *up the steep bank* and *in the savings bank*. All the words except *the* would have similar fans, and her algorithm would "pare down" the ambiguities "by retaining only the spokes that retain ideas which occur in each." For this example, it would retain "OBLIQUITY 220 in 'steep' and 'bank'; whereas it retains as common between 'savings' and 'bank' both of the two areas STORE 632 and TREASURY 799." She went on to discuss methods of handling various exceptions and complications, but all the algorithms use only words and families of words that actually occur in English. They never use abstract or artificial markers, features, or categories. That approach suggests a plausible cognitive theory: From an infant's first words to an adult's level of competence, language learning is a continuous process of building and refining the stock of words, families of words grouped by use in the same contexts, and patterns of connections among the words and families.

Wittgenstein's language games and the related proposals by Cruse, Croft, and Masterman are more realistic models of natural language than the rigid theories of formal semantics. Yet scientists, engineers, and computer programmers routinely produce highly precise language-like structures by disciplined extensions of the methods used for ordinary language. Furthermore, the level of precision needed to write computer programs can be acquired by school children without formal training. A complete theory of language must be able to explain every level of competence from the initial vague stages to the most highly disciplined representations and reasoning methods of science. Different language games may require attention to different details with different granularity, but there is no evidence for a discontinuity in the methods of language generation and understanding.

6. Language Games as a Basis for Semantics

To handle both formal and informal language, Masterman's approach must be extended with links to logic, but in a way that permits arbitrary revisions, changes of perspective, and levels of granularity. Figure 2 illustrates the issues in relating logic, models, and the world. At the right is a theory expressed in the Peirce-Peano notation for logic. In the middle is a formal model shown as a graph in which nodes represent objects and arcs represent relations among those objects. With varying degrees of formality, logicians from Aristotle and the medieval Scholastics to Bolzano, Peirce, Wittgenstein, and Tarski reached a consensus on how to evaluate the denotation of a proposition in terms of a model. But on the left of Figure 2, the mapping of models to the world is an approximation that raises the most contentious issues. As the engineer and statistician George Box (2005) said, "All models are wrong; some are useful."

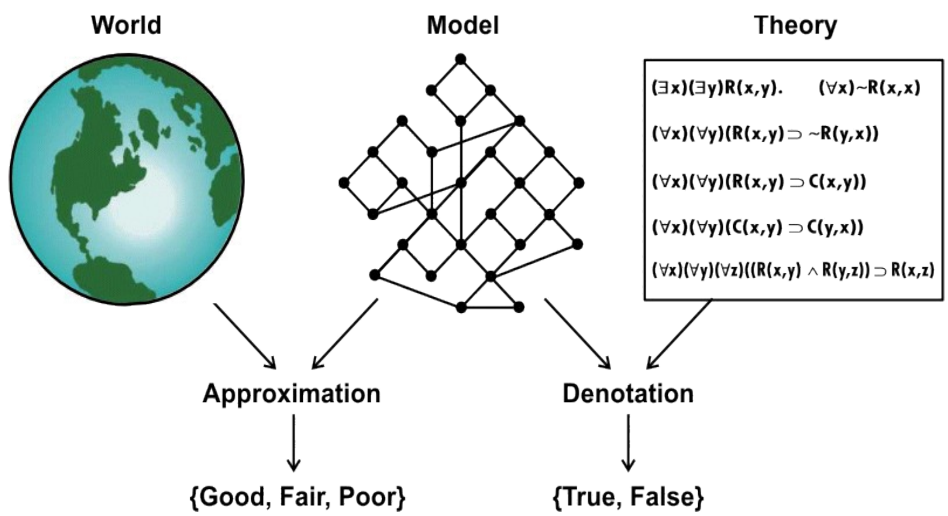

Figure 2: The world, a model, and a theory

The approximate mapping of models to the world is the source of the vagueness that must be addressed in every theory of epistemology, ontology, phenomenology, and philosophy of science. In the *Tractatus*, Wittgenstein assumed an exact mapping from language to logic to models to the world. As he said,

> "The totality of true thoughts is a picture of the world" (3.01). "The picture is a model of reality" (2.12). "The proposition is a picture of reality, for I know the state of affairs presented by it, if I understand the proposition" (4.021). "Reality is determined by the truth or falsity of the proposition; it must therefore be completely described by the proposition" (4.023).

Tarski (1933) was more cautious. He avoided the complexities of natural language and the world by limiting his claims to the relationship between a formal language and a model. Carnap, Kripke, and Montague extended Tarski's approach to modal logic by assuming a multiplicity of models, one of which represents the real world and the others represent possible worlds. Barwise and Perry (1983) avoided a giant model of everything by assuming finite chunks of the world called *situations*. Yet as

Devlin (1991) observed, nobody could state the criteria for selecting significant chunks: “Situations are just that: situations. They are abstract objects introduced so that we can handle issues of context, background, and so on.” In short, situations determine meaning, but there are no criteria for distinguishing a meaningful situation from an arbitrary chunk of space-time.

In his later philosophy, Wittgenstein shifted the focus from abstract mappings between language and the world to the human activities that give meaning to chunks of the world and the language about them. To accommodate language games in a framework that can represent any theory about any model for any purpose, Sowa (2000) proposed an infinite lattice of all possible theories expressible in a given logic. Each theory would represent the rules or axioms of one language game or a family of closely related games. The lattice is a generalization hierarchy, in which the most general theory at the top is true for every possible model; the bottom is the inconsistent theory that is false for every model. Every theory in between is true for a subset of the models of the theories above it and a superset of the models of the theories below it. Figure 3 shows the four basic operators for navigating the lattice: *contraction, expansion, revision,* and *analogy*.

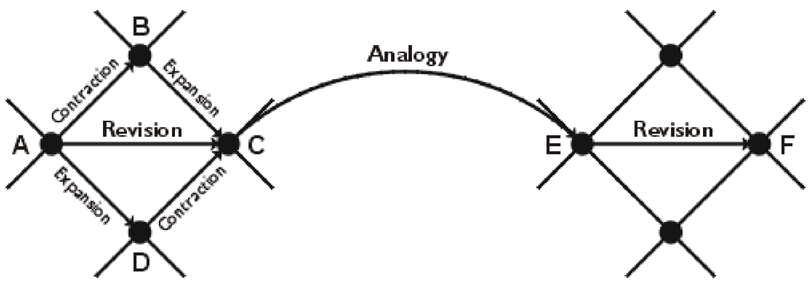

Figure 3: Four operators for navigating the lattice of theories

The operators of contraction and expansion follow the arcs of the lattice, revision makes short hops sideways, and analogy makes long-distance jumps. The first three operators, which delete and add axioms, correspond to the AGM operators for theory revision (Alchourrón et al. 1985). The analogy operator makes longer jumps through the lattice by systematically relabeling the names of types and relations. All methods of nonmonotonic reasoning can be viewed as strategies for walking or jumping through the lattice in order to find a theory that is a suitable approximation to some aspect of the world for some purpose:

- *Induction* is an expansion strategy for increasing the number of provable statements (theorems) while reducing the number of assumptions (axioms).
- *Abduction* is another expansion strategy, which often uses analogy to “guess” or hypothesize a likely theory, whose predictions by deduction are tested against further observations.
- *Default logics* can be considered shorthand descriptions for families of closely related theories. The *supremum* or most specific common generalization of all theories in a family is the classical theory obtained by ignoring all defaults. Other theories in the family are obtained by expanding the supremum with one or more of the defaults.
- *Negation by failure* is a variant of default logic. The supremum is a theory defined by the conjunction of a given set of axioms. Each failure to prove some proposition *p* expands the current theory with the negation *~p*.
- Reasoning methods that use *certainty factors* or *fuzzy values* can be viewed as variants of a default logic in which each proposition has a metalevel measure of its approximation to some aspect of the world. The result of fuzzy reasoning is a theory whose propositions exceed some minimum level of approximation.

These reasoning methods have a common goal: the discovery or construction of an appropriate theory somewhere in the lattice. Combinations of various methods may be applied iteratively to derive theories whose models are better and better approximations to the world.

Figure 4 illustrates a word fan that maps the words of a language to concept types to canonical graphs and to a lattice of theories. The fan on the left of Figure 4 links each word to an open-ended list of *concept types*, each of which corresponds to some area of a thesaurus, as in Masterman's system. The word *bank*, for example, could be linked to types with labels such as `Bank799` or `Bank_Treasury`.

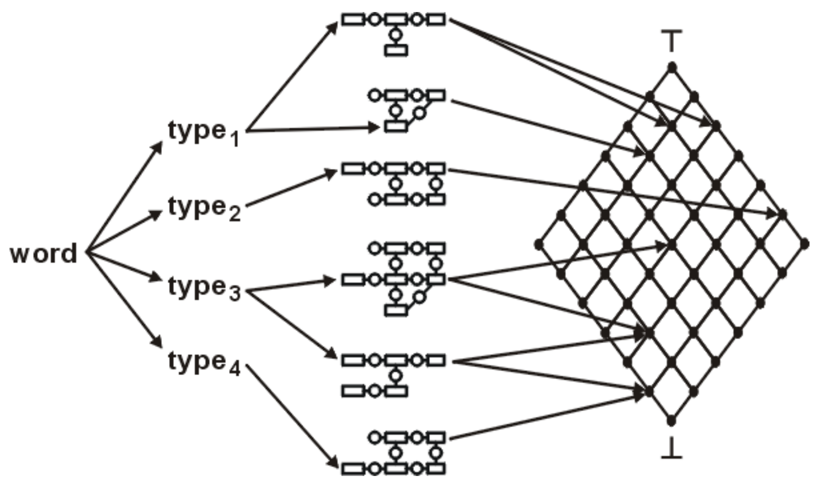

Figure 4: words → types → canonical graphs → lattice of theories

In various language games, those types could be further specialized in subtypes, which would correspond to Cruse's microsenses. When precision is necessary, the lattice enables any theory to be specialized, revised, or refined in order to tighten the constraints or add any amount of detail. In a formal logic, vagueness is not possible, but vagueness in natural language can be represented in two ways: first, the types and theories at the upper levels of the lattice may be underspecified to include a broad range of more specialized language games at lower levels; second, some canonical graphs may lead to more than one theory, and further information may be needed to determine which one is intended.

For this article, canonical graphs are represented by *conceptual graphs* (CGs), a formally defined version of logic that uses the model-theoretic foundation of Common Logic (ISO/IEC 2006). Equivalent operations may be performed with other notations, but graphs support highly structured operations that are computationally more efficient and cognitively more realistic than the rules of inference of predicate calculus (Sowa and Majumdar 2003). Figure 5 illustrates three canonical graphs for the types **`Give`**, **`Easy`**, and **`Eager`**.

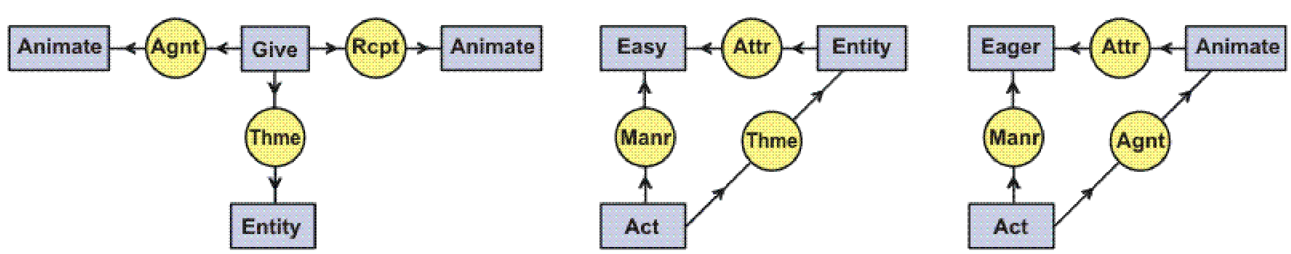

Figure 5: Canonical graphs for the types `Give`, `Easy`, and `Eager`

A canonical graph for a type is a conceptual graph that specifies one of the patterns characteristic of that type. On the left, the canonical graph for **`Give`** represents the same constraints as a typical *case frame* for a verb. It states that the agent **`(Agnt)`** must be **`Animate`**, the recipient **`(Rcpt)`** must be **`Animate`**, and the object **`(Obj)`** may be any **`Entity`**. The canonical graphs for **`Easy`** and **`Eager`**, however, illustrate the advantage of graphs over frames: a graph permits cycles, and the arcs can distinguish the directionality of the relations. Consider the following two sentences:

Bob is easy to please. *Bob is eager to please.*

For both sentences, the concept **`[Person: Bob]`** would be linked via the attribute relation **`(Attr)`** to the concept **`[Easy]`** or **`[Eager]`**, and the act **`[Please]`** would be linked via the manner relation **`(Manr)`** to the same concept. But the canonical graph for `Easy` would make Bob the object of **`Please`**, and the graph for **`Eager`** would make Bob the agent. The first sentence below is acceptable because the object may be any entity, but the constraint that the agent of an act must be animate would make the second unacceptable:

The book is easy to read. ** The book is eager to read.*

Chomsky (1965) used the easy/eager example to argue for different syntactic transformations associated with the two adjectives. But the canonical graphs state semantic constraints that cover a wider range of linguistic phenomena with simpler syntactic rules. A child learning a first language or an adult reading a foreign language can use semantic constraints to interpret sentences with unknown or even ungrammatical syntax. Under Chomsky's hypothesis that syntax is a prerequisite for semantics, such learning is inexplicable.

Canonical graphs with a few concept nodes are adequate to discriminate the general senses of most words, but the canonical graphs for detailed microsenses can become much more complex. The microsenses for the adjective *easy* occur in very different patterns for a book that's easy to read, a person that's easy to please, or a car that's easy to drive. For the verb *give*, a large dictionary lists dozens of senses, and the number of microsenses is enormous. The prototypical act of giving is to hand something to someone, but a large object can be given just by pointing to it and saying "It's yours." When the gift is an action, as in giving a kiss, a kick, or a bath, the canonical graph used to parse the sentence has a few more nodes. But the graphs required to understand the implications of each type of action are far more complex, and they're related to the graphs for taking a bath or stealing a kiss.

The canonical graph for *buy* typically has two acts of giving: money from the buyer to the seller, and goods from the seller to the buyer. But the graphs for specialized microsenses may have far more detail about the buyers, the sellers, the goods sold, and other people, places, and things involved. Buying a computer, for example, can be done by clicking some boxes on a screen and typing the billing and shipping information. That process may trigger a series of international transactions, which can be viewed by going to the UPS web site to check when the computer was airmailed from Hong Kong and delivered to New York. In talking or reasoning about a successful purchase, most of the detail can be ignored, but it may become important if something goes wrong.

7. Language, Logic, and Lebensform

The role of logic in natural language semantics is a controversial issue. Although Montague rejected Chomsky's emphasis on syntax, he adopted Chomsky's distinction between competence and performance, but with semantics at the focus. Instead of an idealized syntax that characterizes the ultimate human competence, Montague (1970) assumed "a theory of truth, of a formal language that I believe may be reasonably regarded as a fragment of ordinary English." But a cognitively realistic theory must also address the question of how that competence is acquired. At age three, Laura correctly used the words *can* and *could* to contrast her own linguistic abilities at different points in time. Presumably that implies a competence for conceiving different contexts, comparing what is possible in each, and expressing her conclusions in English. Yet it seems unlikely that a three-year-old child would have the full logical machinery of Montague's possible worlds.

Linguists and logicians working in Montague's tradition have refined, extended, and restricted his logic in various ways. Fox and Lappin (2005), for example, developed Property Theory with Curry Typing (PTCT) as "a first-order representation language that provides fine-grained intensionality, limited expressive power, and a richly expressive type system." Any such proposal for an ideal formal logic raises some serious issues:

1. Is that formal logic innate? Or does a child learn it in successive stages? As Laura's speech indicates, the semantics for some version of metalanguage and modal logic is acquired very early. But how expressive are those early stages, and how are they learned?
2. Languages such as Pirahã show that an entire community can live successfully without having any native speaker who has achieved the logical sophistication assumed by systems such as Montague's or PTCT. Does that imply that different languages have different kinds of semantic competence? Or that some don't reach the ultimate level of human competence? Or that semantics can be revised and extended indefinitely with no fixed limit?
3. Scientists often invent radically new theories whose mathematical foundations are quite different from any version of formal semantics. When two mathematicians talk about their theories on the telephone, they use the linguistic forms of their native language without the aid of other notations. Does that imply that the formal logic that characterizes their speech must incorporate the semantics of the mathematics they conceived? Does there exist any fixed logic that can characterize everything that is humanly conceivable? Or does Gödel's undecidability theorem rule out that possibility?
4. Did human semantic competence evolve from a more primitive stage around the time of *Homo habilis*, about two million years ago? Or did it spring full-blown into the psyche of Adam and Eve, perhaps 60 thousand years ago? If it didn't evolve, why did the human vocal tract and brain size take a few million years to attain their current forms? If it did evolve, what kinds of intermediate stages could there be?

In the *Tractatus*, Wittgenstein proposed a first-order semantic theory that was far more restricted than Montague's or PTCT. It could not characterize the speech of his pupils in the Austrian village. In the *Philosophical Investigations*, he said that "to imagine a language (eine Sprache vorstellen) is to imagine a form of life (Lebensform)" (§19). Every form of life determines one or more language games, which impose requirements on the expressive power of the associated logic. The various forms of life would include the activities of hunting and gathering by the Pirahã or the so-called "civilized" activities of shopping in a supermarket, reporting a medical diagnosis, and directing traffic around a construction site. Each activity involves constraints imposed by the culture and the environment, which determine the vocabulary, the semantic patterns, and the conventional moves in the corresponding language game.

These considerations suggest that the goal of a fixed formal semantics for all of language is as unrealistic as Hilbert's goal of a fixed foundation for all of mathematics. For many language games, the semantics could be logically simpler than anything required for a general theory of everything. But when new circumstances require changes in the old games or the invention of a totally new game, more complex logical features may be required. Laura's metalevel sentence at age three is considerably more complex than most of her utterances at that age, but it illustrates an important principle: even though most sentences express a rather simple logic, the logical and syntactic complexity increases when someone compares different language games, suggests an innovation in an old game, or proposes a totally new one.

The questions of how language and logic are learned are fundamental to understanding the role of logic in semantics. Frege and Russell, for example, adopted the universal quantifier, negation, and material implication as their three primitives. But those three operators are among the most problematical — logically, linguistically, computationally, and pedagogically. Following is a brief summary of the issues:

- **Existential-conjunctive logic.** Conjunction and the existential quantifier are the two operators that are central to all uses of language. They are the only two that are necessary for observation statements, they are the two most frequently occurring operators in translations from language to logic, and they are needed to represent a child's earliest utterances.
- **Negation.** Words for negation also occur very early in a child's speech, but they raise an enormous number of questions. What aspects of the utterance or the environment do they negate? And do they represent the denial, rejection, absence, or prohibition of those aspects? Many languages use different words or syntax to distinguish different varieties of negation, which must be related to one another.
- **Other logical operators.** Conjunction, negation, and the existential quantifier are sufficient to define all the other operators of first-order logic, but all the problems of negation are inherited by every operator defined in terms of it. Words for many of those operators occur in most natural languages, but they are not the first to be learned, and their semantics is rarely identical to the usual definitions in classical FOL.
- **Commands, statements, and questions.** Imperatives, such as crying for food or attention, precede language, and many of an infant's earliest utterances are refinements of those cries. Without imperatives and interrogatives, declaratives can only paint a static picture of the world. Commands and questions animate that picture and integrate it with the activities that give it meaning.
- **Speech acts.** Peirce, Wittgenstein, Austin, and others studied the use of language, the purpose or intention of any particular statement, and its role in relation to the speaker, listener, discourse, environment, and accompanying activity. Without considering the use, it is impossible for anyone to understand an infant's utterances and often misleading to try to understand an adult's.
- **Context.** Most versions of logic are deliberately designed to have a context-free syntax, but almost all aspects of natural language are context sensitive. Although the word *context* has multiple senses, just the basic definition as a chunk of text is sufficient to raise the questions: How are the contents of one chunk of text related to other chunks, to the environment, to the participants in the discourse, and to the goals of the participants?
- **Metalanguage.** From infancy, children are surrounded by language about language, which they imitate successfully by their third year: praise, blame, corrections, prompts, explanations, definitions, and examples of how language maps to things, activities, and people, including themselves. All the tenses and modalities of verbs are metalinguistic commentary, which can be defined by language about language or logic about logic (Sowa 2003).
- **Propositions.** Some metalanguage is about syntax and vocabulary, but much of it is about language-independent propositions. Many logicians avoid the notion of proposition by talking only about sentences, but that approach ignores the fact that people find it easier to remember what was said than to remember how it was said. Other logicians identify a proposition with the

set of possible worlds in which it is true, but that definition is much too coarse grained. It cannot distinguish 2+2=4 from Fermat's last theorem.

- **Fuzziness.** Hedges and "fuzzy" qualifiers such as *almost* or *nearly* have spawned a variety of fuzzy or multivalued logics with a range of truth values or certainty factors between 1 for true and 0 for false. But many logicians have pointed out the problems with interpreting those numbers as truth values. A more nuanced approach should observe the distinction in Figure 2: truth values are metalevel commentary about the mapping of a sentence to a model; fuzzy values estimate the adequacy of the model as an approximation to the world for the purpose of a given language game.

Conventional model theory, by itself, is insufficient to accommodate all these aspects of language in a cognitively realistic formalism. Although Wittgenstein contributed to that paradigm, he recognized its limitations and proposed language games as an alternative. The challenge is to formalize language games and integrate them with related research in cognitive science.

Some promising techniques published decades ago were ignored because they did not fit the popular paradigms. Among them are the *surface models* by Hintikka (1973). Like situations, surface models are finite. But unlike situations, which are considered chunks of the world, surface models are constructed as approximations to the world, as in Figure 2. Instead of trying to define criteria for meaningful situations, Hintikka proposed a method for constructing surface models that represent the individuals and relations explicitly mentioned in the discourse. In that same book, Dunn (1973) published an alternative semantics for modal logics based on sets of *laws* and *facts*. Each possible world *w* is replaced by a pair of sets (**M**,**L**), in which **M** consists of all facts that are true in *w* and **L** consists of the laws of *w* — the subset of facts that are necessarily true. Dunn showed that this construction is isomorphic to Kripke semantics, it avoids the dubious ontology of possible worlds, and it treats accessibility as a derived relation instead of a primitive. Sowa (2003) showed that Dunn's semantics simplifies the computational and the theoretical methods by treating multimodal reasoning as metalevel reasoning about the choice of laws. These techniques can be combined with the lattice of theories to formalize language games:

- For each type of language game *g*, define a set **L** of propositions as the laws, rules, or axioms of a theory that characterizes any game of type *g*.
- During the play of a game of type *g*, construct a surface model that is derived from the facts that are consistent with **L** and known or assumed to be true as a result of statements during the play.
- Specialized theories at lower levels of the lattice represent the axioms of possible games, and generalizations higher in the lattice represent the axioms common to a family of games.
- Since any game may be associated with extralinguistic activity, some observable facts about individuals, states, and events may be incorporated in the surface model without being explicitly mentioned in language.
- Any fact that is inconsistent with the current game triggers theory revision operations that move through the lattice to find a theory of a related game consistent with that fact.

This approach retains the power and precision of formal methods within a dynamically extensible or negotiable framework. The construction of a surface model need not be monotonically increasing, since various statements, observations, and objections may trigger revisions — either to the surface model or to the laws of the language game that governs its construction. The result of a successful dialog or negotiation is a surface model that is consistent with the axioms of some theory in the lattice and the facts agreed or observed during the discourse. But not all discourse reaches a settled conclusion. Some participants may refuse to accept some statements about the laws and facts, or they may take action to change them.

The most promising and most neglected work is Peirce's research on semiotics and its relationships to both logic and language (Sowa 2006). Many aspects that Peirce discovered, anticipated, or developed in detail are usually associated with other philosophers and logicians:

- Tarski: model theory and metalanguage.
- Davidson: event semantics.
- Austin: speech acts.
- Grice: conversational implicatures.
- Perry: the essential indexical.
- Kamp: nested contexts for discourse representation structures.
- Carnap, Kripke, Montague: possible worlds.

Some of the more recent developments have gone into much greater detail than Peirce had. But Peirce demonstrated that these and other aspects of language are part of a unified vision. Furthermore, Peirce's "left-handed brain," as he called it, often put old ideas in surprisingly new perspectives.

As an example, Peirce observed that a proposition corresponds to "an entire collection of equivalent propositions with their partial interpretants" (CP 5.569). To formalize that insight, a proposition may be defined as an equivalence class of sentences in some language **L** under some meaning-preserving translation (MPT) defined over the sentences of **L**. An MPT is then defined as any function *f* over sentences of **L** that satisfies four constraints: invertible, truth preserving, vocabulary preserving, and structure preserving. If *f* satisfies only the first two constraints, the equivalence classes are much too big: each would consist of all sentences that are true in a given set of possible worlds. Furthermore, that function is not efficiently computable because proving that 2+2=4 is in the same equivalence class as Fermat's last theorem took three centuries of research by the best mathematicians in the world.

If the constraints on vocabulary and structure are too strong, the MPT *f* becomes the identity function, which is trivially computable, but it leaves only one sentence type in each class. By imposing reasonable constraints on vocabulary and structure, Sowa (2000) defined several MPTs that are cognitively realistic and computable in polynomial or even linear time. These functions can be specified in just a few lines, the translations can be learned in pedagogically simple steps, and the method is sufficiently flexible to allow different options of MPTs for different language games. By contrast, the proposal of Curry typing (Fox and Lappin 2005) is a fixed, rigid system that takes 40 pages to specify and makes no provision for learnability.

In summary, language games can be formalized in an open-ended framework that can accommodate any use of language for any purpose. At one extreme are the versions of mathematics and logic with specialized ontologies designed for science and engineering. At the other extreme are the vague ideas and insights whose consequences are not well understood. In between are the discussions, negotiations, compromises, and analyses that are necessary to translate a vague idea to a precise plan or to revise the plan as circumstances change. To adopt this approach requires a major paradigm shift in formal semantics. It doesn't reject logic, but it applies logic to a broader range of problems with a greater sensitivity to the way language is actually used by people at every stage of life.

References

Alchourrón, Carlos, Peter Gärdenfors, & David Makinson (1985) "On the logic of theory change: partial meet contraction and revision functions," *Journal of Symbolic Logic* **50:2**, 510-530.

Box, George E. P., J. Stuart Hunter, & William G. Hunter (2005) *Statistics for Experimenters: Design,* *Innovation, and Discovery*, 2nd Edition, Wiley-Interscience, New York.

Chomsky, Noam (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge, MA.

Croft, William (2001) *Radical Construction Grammar: Syntactic Theory in Typological Perspective*, Oxford University Press, Oxford.

Croft, William, & D. Alan Cruse (2004) *Cognitive linguistics*, Cambridge University Press, Cambridge.

Cruse, D. Alan (2000) “Aspects of the micro-structure of word meanings,” in Y. Ravin & C. Leacock, eds. (2000) *Polysemy: Theoretical and Computational Approaches*, Oxford University Press, Oxford, pp. 30-51.

Devlin, Keith (1991) “Situations as mathematical abstractions,” in J. Barwise, J. M. Mark Gawron, G. Plotkin, & S. Tutiya, eds., *Situation Theory and its Applications*, CSLI, Stanford, CA, pp. 25-39.

Dunn, J. Michael (1973) “A truth value semantics for modal logic,” in H. Leblanc, ed., *Truth, Syntax and Modality*, North-Holland, Amsterdam, pp. 87-100.

Everett, Daniel L. (2005) “Cultural Constraints on Grammar and Cognition in Pirahã,” *Current Anthropology* **46:4**, pp. 621-646.

Fox, Chris, & Shalom Lappin (2005) *Foundations of Intensional Semantics*, Blackwell, Oxford.

Gärdenfors, Peter (2000) *Conceptual Spaces: The Geometry of Thought*, MIT Press, Cambridge, MA.

Hattiangadi, Jagdish N. (1987) *How is Language Possible? Philosophical Reflections on the Evolution of Language and Knowledge*, Open Court, La Salle, IL.

Hintikka, Jaakko (1973) “Surface semantics: definition and its motivation,” in H. Leblanc, ed., *Truth, Syntax and Modality*, North-Holland, Amsterdam, 128-147.

ISO/IEC (2006) *Common Logic: A Framework for a Family of Logic-Based Languages*, Final Committee Draft, available at http://common-logic.org

Kittredge, Richard, & John Lehrberger, eds. (1982) *Sublanguage: Studies of Language in Restricted Semantic Domains*, de Gruyter, New York.

Limber, John (1973) The genesis of complex sentences, in T. Moore, ed., *Cognitive Development and the Acquisition of Language*, Academic Press, New York, 169-186.

Majumdar, Arun K., John F. Sowa, & Paul Tarau (2007) “Graph-based algorithms for intelligent knowledge systems,” to appear in A. Nayak & I. Stojmenovic, eds., *Handbook of Applied Algorithms*, Wiley & Sons, New York.

Masterman, Margaret (2005) *Language, Cohesion and Form*, edited by Yorick Wilks, Cambridge University Press, Cambridge.

Montague, Richard (1970) “English as a formal language,” reprinted in Montague (1974), *Formal Philosophy*, Yale University Press, New Haven, pp. 188-221.

Peirce, Charles Sanders (CP) *Collected Papers of C. S. Peirce*, ed. by C. Hartshorne, P. Weiss, & A. Burks, 8 vols., Harvard University Press, Cambridge, MA, 1931-1958.

Perlis, Alan J. (1982) “Epigrams in Programming,” *SIGPLAN Notices*, Sept. 1982, ACM. Available at http://www.cs.yale.edu/homes/perlis-alan/quotes.html

Simons, Peter (1992) *Philosophy and Logic in Central Europe from Bolzano to Tarski*, Kluwer Academic Publishers, Dordrecht.

Sowa, John F. (1984) *Conceptual Structures: Information Processing in Mind and Machine*, Addison-

Wesley, Reading, MA.

Sowa, John F. (2000) *Knowledge Representation: Logical, Philosophical, and Computational Foundations*, Brooks/Cole Publishing Co., Pacific Grove, CA.

Sowa, John F. (2003) “Laws, facts, and contexts: Foundations for multimodal reasoning,” in *Knowledge Contributors*, edited by V. F. Hendricks, K. F. Jørgensen, and S. A. Pedersen, Kluwer Academic Publishers, Dordrecht, pp. 145-184. http://www.jfsowa.com/pubs/laws.htm

Sowa, John F. (2006) “Peirce’s contributions to the 21st Century,” Published in H. Schärfe, P. Hitzler, & P. Øhrstrøm, eds., *Conceptual Structures: Inspiration and Application*, LNAI 4068, Springer, Berlin, pp. 54-69.

Sowa, John F., & Arun K. Majumdar (2003) “Analogical reasoning,” in A. de Moor, W. Lex, & B. Ganter, eds., *Conceptual Structures for Knowledge Creation and Communication*, LNAI 2746, Springer-Verlag, Berlin, pp. 16-36. http://www.jfsowa.com/pubs/analog.htm

Steiner, George (1975) *After Babel: Aspects of Language and Translation*, Oxford University Press, Oxford, third edition 1998.

Tarski, Alfred (1933) “Der Wahrheitsbegriff in den formalisierten Sprachen,” trans. as “The concept of truth in formalized languages,” in *Logic, Semantics, Metamathematics*, Second edition, Hackett Publishing Co., Indianapolis, pp. 152-278.

Winograd, Terry (1972) *Understanding Natural Language*, Academic Press, New York.

Winograd, Terry, & Fernando Flores (1986) *Understanding Computers and Cognition*, Ablex, Norwood, NJ.

Wittgenstein, Ludwig (1922) *Tractatus Logico-Philosophicus*, Routledge & Kegan Paul, London.

Wittgenstein, Ludwig (1953) *Philosophical Investigations*, Blackwell, Oxford.

Wittgenstein, Ludwig (1958) *The Blue and Brown Books*, Blackwell, Oxford.

Wigner, Eugene (1960) “The unreasonable effectiveness of mathematics in the natural sciences”, *Communications in Pure and Applied Mathematics* **13:1**.

Woods, William A. (1968) “Procedural semantics for a question-answering machine,” *AFIPS Conference Proc.*, 1968 FJCC, 457-471.

Woods, William A., R. M. Kaplan, & B. L. Nash-Webber (1972) *The LUNAR Sciences Natural Language System*, Final Report, NTIS N72-28984.