Book Reviews

Language, Cohesion and Form

Margaret Masterman (1910–1986)
(Edited by Yorick Wilks, University of Sheffield)

Cambridge University Press (Studies in natural language processing, edited by Steven Bird and Branimir Boguraev), 2005, x+312 pp; hardbound, ISBN 0-521-45489-1, $95.00, £55.00; eBook, ISBN 0-511-13318-9, $76.00

Reviewed by
John F. Sowa
VivoMind Intelligence, Inc.

Margaret Masterman was one of six students in Wittgenstein's course of 1933–1934 whose notes were compiled as *The Blue Book* (Wittgenstein 1958, which is the first publication after the mimeographed copies that were circulated informally). In the late 1950s, she founded the Cambridge Language Research Unit (CLRU) as a discussion group, which evolved into one of the pioneering centers of research in computational linguistics. In her will, she requested that Yorick Wilks edit a collection of her papers for publication. The result is this book, which is important for its historical perspective on the development of computational linguistics. It consists of 11 of her reports and articles from the late 1950s to 1980 and includes a 17-page introduction and commentary by Wilks. Karen Spärck Jones also wrote a commentary on one of the papers she had coauthored.

As a student of Wittgenstein, Masterman was also deeply concerned about the foundations of theoretical linguistics. Around the same time that Chomsky was developing his syntactic theories and Montague was advocating a logic-based alternative, she was proposing a "neo-Wittgensteinian" view, whose organizing principle was a thesaurus of words classified according to the "language games" in which they are used. Although no single paper in the book formulates a succinct summary that could be called a theory, the following principles are discussed throughout:

- 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" (page 200).
- Recognition that ambiguity is a consequence of the flexibility and extensibility of natural language and not a defect that can be eliminated by switching to a purified language of logic.
- 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 in the thesaurus to the area in which it occurs, and dynamically generated combinations of fans for 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

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/coli.2006.32.4.551 by guest on 28 March 2024

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 in every field, especially in the most advanced areas of science.

Unlike the a priori formalisms of Chomsky or Montague, this approach is based on data about actual language use. In the commentary, Wilks noted that Masterman's work contained "the germ of what was later to be called EBMT or example-based translation (Nagao 1989), which is now perhaps the most productive current approach to MT world-wide, and I have heard Professor Nagao refer to [her] in this connection in a lecture" (page 279).

As a whole, the book presents a cognitive view of language that has strong similarities to the *Cognitive Linguistics* of Croft and Cruse (2004). Croft's *radical construction grammar*, Cruse's *dynamic construal of meaning*, and Lakoff and Johnson's (1980) work on metaphor are compatible with and to some extent anticipated in Masterman's papers. The multiplicities of context-dependent word senses discussed in the first paper of the book could be aptly characterized by the term *microsense*, which was coined by Cruse (2000). Although most of the papers are forty years old or more, the goal of implementing the ideas in a computable form has forced a greater attention to detail and precision than is found in some of the more recent work on cognitive linguistics.

The age and the origin of most of the papers as unpublished memos is evident in their rather disorganized structure, but the book contains many intriguing insights that still seem fresh today. Among them are her penetrating criticisms of Chomsky's fixation on syntax:

My quarrel with [the Chomsky school] is not that they are abstracting from the facts. How could it be? For I myself in this paper am proposing a far more drastic abstraction from the facts. It is that they are abstracting from the wrong facts because they are abstracting from the syntactic facts, that is, from that very superficial and highly redundant part of language that children, aphasics, people in a hurry, and colloquial speakers, quite rightly, drop. (page 266)

As an alternative, she discussed the writings of the phoneticist Peter Guberina (1954), who had worked in a school for the deaf:

A large part of Guberina's daily life is spent in developing electronic techniques for helping the deaf to speak. This means that, for him, what is being talked about — that is, the actual subject of any piece of discourse, and the linguistic elements that carry it — is vastly more important than what is said about it. If the deaf man can once pick up the subject of conversation, three-quarters of this problem is solved, even if he cannot hear all that is said about it. If, on the other hand, he clearly hears some one thing that is clearly said about some basic subject of discourse, while the actual subject of discourse remains unknown to him, very little of the deaf man's problem is solved; he has only heard one thing. (page 228)

In summary, she said that "human communication consists of patterns of semantic interactions between ascertainably cognate subjects of discourse." By *cognate subjects*, she meant ones that originate from the same or similar language games and are grouped

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/coli.2006.32.4.551 by guest on 28 March 2024

in the same area of a thesaurus. The semantic patterns led to the templates of Wilks's own theory of *preference semantics*, and they are closely related to the chunks, frames, scripts, and schemata of other systems.

We should be grateful to Wilks and the publisher for making these papers available, but the cost of the book and the limited selection underscores the difficulty of getting access to research from the years before the World Wide Web. The eleven papers are a fraction of Masterman's more than 80 published and unpublished reports, memos, and manuscripts, not to mention the more than 200 CLRU memos by other authors. As Wilks noted, there was a great deal of duplication among the manuscripts, but even for those that were selected, much of the lengthy appendices, quotations, and diagrams had to be omitted. One of the papers that was not included (Masterman 1961) contained the first recorded use of the term *semantic network*. The old material is valuable for clarifying the historical record and for making available important, but long-forgotten insights.

References

Croft, William and D. Alan Cruse. 2004. *Cognitive Linguistics.* Cambridge University Press.

Cruse, D. Alan. 2000. Aspects of the micro-structure of word meanings. In Y. Ravin and C. Leacock, editors, *Polysemy: Theoretical and Computational Approaches*, Oxford University Press, pages 30–51.

Guberina, Peter. 1954. *Valeur Logique et Valeur Stylistique des Propositions Complexes*. Zagreb: Editions Epoha.

Lakoff, George and Mark Johnson. 1980. *Metaphors We Live By.* University of Chicago Press.

Masterman, Margaret. 1961. Semantic message detection for machine translation, using an interlingua. In *Proceedings of the First International Conference on Machine Translation of Languages and Applied Language Analysis*, pages 438–475, HMSO, London.

Nagao, Makoto. 1989. *Machine Translation: How Far Can It Go?* Oxford University Press.

Wittgenstein, Ludwig. 1958. *Preliminary Studies for the "Philosophical Investigations": Generally known as The Blue and Brown Books.* Oxford: Blackwell.

John F. Sowa is the president and cofounder of VivoMind Intelligence, Inc. He had earlier retired from IBM after 30 years of working on R&D projects in artificial intelligence, computational linguistics, and related areas of computer science. He is also a fellow of the American Association for Artificial Intelligence. Sowa's e-mail address is sowa@vivomind.com.

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/coli.2006.32.4.551 by guest on 28 March 2024

Downloaded from http://www.mitpressjournals.org/doi/pdf/10.1162/coli.2006.32.4.551 by guest on 28 March 2024