Wittgenstein in the Machine

Lydia H. Liu

Computing is too important to be left to men.

<div align="right">—Karen Spärck Jones</div>

Can a machine have toothache?

There are some questions we should never put to Siri and other talking chatbots, or we will come away feeling underwhelmed by the machine's much-touted intelligence. This could be one of them. The machine will not understand the question—at least, not in the foreseeable future—although humans would not fare much better trying to make sense of it either. The question bears the hallmarks of philosophical surrealism that identify the author as Ludwig Wittgenstein.[1] It is as if the philosopher had anticipated Alan Turing's classic provocation "Can machines think?" and decided to parody him.[2]

Unlike the Turing question, Wittgenstein's parody *avant la lettre* requires no reply, has elicited none, and leads only to further questions: What is the point of his parody? Is it to convey his misgivings about the claims of machine cognition? If that were true, why single out "toothache"—something machines can do happily without—when he could have evoked nobler human aptitudes as do most anthropocentric critics when

1. See Ludwig Wittgenstein, *Preliminary Studies for the "Philosophical Investigations": Generally Known as The Blue and Brown Books* (New York, 1969), p. 16; hereafter abbreviated *PS*.

2. Alan Turing, "Computing Machinery and Intelligence," in Turing et al., *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life, plus the Secrets of Enigma*, ed. B. Jack Copeland (New York, 2004), p. 441.

they engage AI practitioners in contentious debates. Some of us might be tempted to read in Wittgenstein's question an endorsement of affect, emotion, and noncognitive behavior, holding onto the last preserve of human/animal identities in the face of the imminent encroachment of AI affective simulations. Unfortunately, such a straight-faced reading would stretch the sense of *toothache* beyond the word itself until we lose the parody altogether.

When Wittgenstein dictated his lectures in 1933–1934 to his students at the University of Cambridge, where the question first came up, machine cognition and machine affect seemed rather remote from his mind. The sentence "Could a machine think?" struck him as flawed because it exemplified what he called the misused analogy by human speakers (*PS*, p. 16). To allow the analogy to take hold is to grant as much *sense* to the parodic sentence "Can a machine have toothache?" as it is to the proposition of a thinking machine (*PS*, p. 16). Wittgenstein goes on to say: "The trouble is rather that the sentence, 'A machine thinks (perceives, wishes)': seems somehow nonsensical. It is as though we had asked 'Has the number 3 a colour?'" (*PS*, p. 47). Which is to say that, prior to advancing a possible argument against the intellectual claims of the computing machine, one must interrogate and critique the *sense* of the original proposition first, and this is where his parody comes in.[3]

In a surprising turn of events, that critique anticipated Wittgenstein's subsequent encounter and arguments with Turing in 1939 when the latter showed up in his class. Their open confrontation has prompted Stuart Shanker to speculate that "the Turing Test represents Turing's opposition to Wittgenstein's critique, using a Wittgenstein-like argument."[4] Speculations aside, it is reasonable to observe in retrospect that the fateful encounter between Wittgenstein and Turing in 1939 was a mere prelude to the

3. In *Philosophical Investigations*, Wittgenstein repeats the point and considers what pertains between *saying* and *thinking* when we project the word *think* elsewhere, to dolls, ghosts, and machines. This is followed by another parody: "The chair is thinking to itself . . ." (Wittgenstein, *Philosophical Investigations*, trans. G. E. M. Anscombe, P. M. S. Hacker, and Joachim Schulte [Malden, Mass., 2009], p. 121e).

4. Stuart Shanker, *Wittgenstein's Remarks on the Foundations of AI* (New York, 1998), p. 2.

LYDIA H. LIU is the Wun Tsun Tam Professor in the Humanities at Columbia University and Director of the Institute for Comparative Literature and Society. She is the author of *The Freudian Robot: Digital Media and the Future of the Unconscious* (2010); *The Clash of Empires: The Invention of China in Modern World Making* (2004); and *Translingual Practice: Literature, National Culture, and Translated Modernity* (1995).

full-blown postwar debates about artificial intelligence.[5] Philosophers who stand in opposition to the exaggerated claims of AI inventions believe that they have found a powerful ally in Wittgenstein. Hubert Dreyfus, for example, has repeatedly cautioned us about what the computer cannot do, citing human "fringe consciousness," "ambiguity tolerance," and so on, and he draws on Wittgenstein to shore up the position.[6]

   It seems unthinkable, therefore, that Wittgenstein or the philosophy his name stands for has anything to do with the computing machine itself, much less with the AI machine. But this is exactly what I intend to demonstrate in my essay, and I am going to present my evidence by calling attention to a parallel postwar development surrounding Wittgenstein that is just as important as the familiar critical stance. This alternative deep history began to emerge as I was examining a counterintuitive situation at the initial stage of my research, namely, Why did Wittgenstein's profound doubts and open disagreements with Turing fail to deter leading AI precursors and practitioners from claiming him as one of their own?

   The news of AI researchers' longtime engagement with Wittgenstein has been slow to arrive. The truth is that Wittgenstein's philosophy of language is so closely bound up with the semantic networks of the computer from the mid-1950s down to the present that we can no longer turn a blind eye to its embodiment in the AI machine. But what does it mean for AI practitioners to engage with Wittgenstein? Let me illustrate it by giving you a contemporary example. John F. Sowa, a prominent AI scientist who invented conceptual graphs for databases in the 1970s, has attributed the distinction he introduces between lexical structure and conceptual structure to Wittgenstein, who taught him that "the ambiguities and complexities of language result from its use in novel situations with novel ways of relating words to objects" and, therefore, "the ultimate source of ambiguity is not the structure of language, but the complexity and variability of the world itself."[7] We need not take Sowa's interpretation as a faithful representation of Wittgenstein's own ideas to heed the fact that the computer scientist is grappling with the philosophical implications of his work *in Wittgensteinian terms* as he tries to resolve the technical difficulties of word and concept entanglement in the computer—that is, units, levels, and frames of knowledge representation. The larger implication of Sowa's theory is not something that I can elaborate on in my article. What I want to emphasize is that his machine is Wittgensteinian in the sense that the

   5. See Hubert Dreyfus's defense of Wittgenstein against Turing in Hubert L. Dreyfus, *What Computers Can't Do: A Critique of Artificial Reason* (New York, 1972), pp. 104–05.
   6. Ibid., pp. 100, 107.
   7. John F. Sowa, "Lexical Structures and Conceptual Structures," in *Semantics and the Lexicon*, ed. James Pustejovsky (Boston, 1993), p. 249.

conceptual graphs describe the meaning of data in accordance to the user's view even though these are also associated with procedures that access the data according to the machine's view.[8] That is to say, the multiple human uses of language determine the multiplicity of meanings that the conceptual graphs must accommodate in the machine. Sowa is not alone nor the first in this endeavor.

Which brings me to British AI scientist Yorick Wilks, who has always insisted that AI concerns were present in Wittgenstein's later philosophy.[9] What Wilks seems to be claiming goes deeper and further than the instrumental application of Wittgenstein's ideas to the computer as some are inclined to regard Sowa's research. Wilks and Sowa both work in the key subfields of AI research, notably Natural Language Processing (NLP, formerly known as Computational Linguistics), Machine Translation (MT), and Information Retrieval (IR). Researchers in these fields find themselves in the unenviable situation of having to grapple with the challenges of ordinary language and confront its unique difficulties and infinite perplexities. Having struggled with the same difficulties, Wittgenstein would have been in a perfect position to dialogue with the newcomers. One of the world's pioneering teams of AI researchers that I discuss below has not hesitated to trace the philosophical roots of their work to Wittgenstein.

I wonder sometimes if the AI scientist's Wittgenstein is the same thinker as the philosopher's Wittgenstein. Dreyfus would almost certainly say no, whereas Sowa, Wilks, and others would probably say yes. There is no doubt that Wittgenstein's philosophy can be open to all sorts of interpretations—after all, the philosopher repudiated his own *Tractatus Logico-Philosophicus* in later life—and he has been read differently by linguists, philosophers, and others. Still, one must account for the fact that his work has engaged the attention of AI researchers not only in the 1970s down to the present but also from the beginnings of computational research. In particular, Wittgenstein inspired a group of researchers called Cambridge Language Research Unit (CLRU) in Britain to launch one of the first programs in machine translation, information retrieval, mechanical abstracting, and so on in the 1950s, all of which are now claimed for AI and cognitive science.

It is in the philosophy of CLRU founder Margaret Masterman (1910–1986) that we will find our first clue as to how Wittgenstein or his later philosophy got into the AI machine, both literally and metaphorically. Like

8. See Sowa, "Conceptual Graphs for a Data Base Interface," *IBM Journal of Research and Development* 20 (July 1976): 336–57.

9. See Yorick Wilks, "Philosophy of Language," in *Computational Semantics: An Introduction to Artificial Intelligence and Natural Language Comprehension*, ed. Eugene Charniak and Wilks (New York, 1981), pp. 205–33.

all such stories, there is an embedded story of gender hidden in plain sight, as suggested by the epigram from eminent AI scientist Karen Spärck Jones: "Computing is too important to be left to men."[10] The story of a woman philosopher of AI has been waiting to be told, as long as one does not to suppress women on purpose or subconsciously. It is a deeply human story about how the pioneers in AI research struggled to make sense of the infinitely entangled web of words and concepts in the language machine. The reciprocal illumination of humans and machines in the course of that struggle should give us some inkling of what is possible in the philosophy of ordinary language. One glimmer of hope comes from the direction of how—not whether—Western philosophy will move beyond its self-imposed myopia and open up to other possibilities; I mean genuine possibilities beyond the ethnocentric imagination of analytic philosophy or continental philosophy that has heretofore dominated our conception of language, logic, writing, and media technology. Hence, the urgency and importance of learning from post-Wittgensteinian philosophy and carrying its work forward.

The Woman Philosopher Who Pioneered the AI Machine

The posthumous publication of Wittgenstein's *Philosophical Investigations* (1953) and *The Blue and Brown Books* (1958) coincided with the emergence of cybernetics and communication technologies that sought to redefine the terms of philosophical inquiries about language, meaning, and the mind. It was a time when Noam Chomsky was developing his theory of syntactic structure and transformational generative grammar at MIT, and, at the same time, Margaret Masterman was building her thesaurus model and semantic algorithms with the young researchers that she recruited to CLRU in Cambridge, UK. Unlike Chomsky, Masterman's point of departure was not syntactic rules but *semantic nets* or *networks* of words or semantic patterns that form the basis of the machine representation of knowledge and its processing of natural language. In the history of AI research, it is well known that her work was closely related to the development of the chunks, frames, scripts, and schemata of AI systems and inspired Wilks's pathbreaking theory of preference semantics. Speaking of Masterman's contributions to machine translation, Wilks himself notes that her innovation in the 1950s contains "the germ of what later was to be called EBMT or example-based translation . . . , which is now perhaps the most productive current approach to MT world-wide."[11] In hindsight—like

10. Quoted in Nellie Bowles, "Overlooked No More: Karen Sparck Jones, Who Established the Basis for Search Engines," *New York Times*, 2 Jan. 2019, www.nytimes.com/2019/01/02/obitu aries/karen-sparck-jones-overlooked.html?searchResultPosition=1

11. Wilks, commentary to Margaret Masterman, "Semantic Algorithms," in *Language, Cohesion and Form*, ed. Wilks (New York, 2005), p. 279; hereafter abbreviated "S."

the advances in statistical models of MT and the decline of syntax analysis since the 1990s—her work proved extraordinarily prescient. Her thesaurus method and semantic nets, in particular, put Masterman many decades ahead of her time.

The profound and irreconcilable chasm between Masterman's model of language and that of Chomsky's is something she insisted upon throughout the 1950s–1960s, and she never missed an opportunity to reiterate their philosophical differences. Her main objection to the prevailing theory of syntactic structure is that Chomsky's syntactic rules are modeled on logical calculus, not on natural languages that are flexible, rich, ambiguous, metaphorical, and infinitely extensible. Like other rules derived from the calculus, syntactic rules subtract their linguistic facts "from that very superficial and highly redundant part of language that children, aphasics, people in a hurry, and colloquial speakers always, quite rightly, drop" ("S," p. 266). Masterman contends that the ambiguities and indeterminate meanings in natural languages are not a defect to be overcome by substituting a purified language of logical calculus. On the contrary, the key to understanding natural language and, consequently, its adequate coding on the computer, must be sought in semantic networks that alone are capable of handling the multiplicity and indeterminacy of word meanings. What this means is that one must focus on *data in actual language use*, not on what the MT linguists were practicing at that time: sentence parsing.

If Masterman's defense of ordinary language on the MT machine bears uncanny resemblance to Wittgenstein's ideas, it need not surprise us, as there had been an intellectual bond between them, one that had originated in the interwar years when she first encountered him in a class at the University of Cambridge. That bond underwent an unexpected transformation to evolve into what I term post-Wittgensteinian philosophy after World War II. Masterman vigorously pursued her philosophical work in both the traditional venues of philosophy—debates, seminars, workshops, journals, and conference proceedings—and in its algorithmic embodiment by the computing machine. It cannot be fortuitous that the woman who led one of the world's leading research centers on computation research to lay the foundation of essential AI technologies was a philosopher, not a computer scientist.

Who was Margaret Masterman?

Not much biographical information is available to give us the full life story of this extraordinary woman except what we can glean from an account by Wilks.[12] There are also some scattered sketches and summaries

---

12. See Wilks, "Margaret Masterman," in *Early Years in Machine Translation: Memoirs and Biographies of Pioneers*, ed. W. John Hutchins (Philadelphia, 2000), pp. 279–97.

across journals and academic books, such as a fascinating piece by Kwame Anthony Appiah. In his reminiscences about the mid-1970s, Appiah recounts how he wandered into Masterman's circle as a college student. He writes:

> My main intellectual mentor at Cambridge wasn't a person: it was a commune. Its members called themselves the Epiphany Philosophers, or E.P.'s. The central figures Dorothy Emmet, Margaret Masterman and Masterman's husband, Richard Braithwaite lived together in a large house, round the corner from a ramshackle shed they also owned, home to the Cambridge Language Research Unit, or C.L.R.U., which Masterman ran. What they had in common was not just philosophy and a taste for abbreviations. . . . The C.L.R.U. was past its prime, though through its portals had passed some of the people, like Roger Needham and Karen Spärck Jones, who helped lay the intellectual foundations of modern computing.
>
> The E.P.'s were brilliant, argumentative, generous and, often, quite dotty. . . . It was the E.P.'s who introduced me to the possibility of philosophy, not just as an academic subject but as a way of life. Philosophy for them was mixed up with friendship and brisk walks in the Norfolk countryside and drinking cheap (and even, occasionally, expensive) wine; it required openness to physics, linguistics, theology, parapsychology. Nothing human or otherwise was alien to them. For a 19-year-old, it was completely exhilarating.[13]

This charming vignette of a small but open intellectual community centering around the three philosophers communicates a tangible sense of how members of CLRU lived and worked together twenty years after its founding. Roger Needham (1935–2003) and Spärck Jones (1935–2007) had been Masterman's longtime collaborators at CLRU.[14] In his study of William Empson in light of vector semantics, Michael Gavin indicates that, together with Spärck Jones, "Masterman developed the first computer-based thesaurus, drawn from *Roget's*, for modeling word meaning; words, Masterman believed, distributed meaning through a corpus in a lattice-shaped network."[15] This is true, and it remains to be seen how much of Masterman's philosophical work went into the computer modeling of word meanings.

---

13. Kwame Anthony Appiah, "The Epiphany Philosophers," *New York Times Magazine*, 19 Sept. 2008, www.nytimes.com/2008/09/21/magazine/21jolley-side.html?searchResultPosition=1

14. The core members of CLRU researchers were not more than ten people. Three of its alumni Martin Kay, Spärck Jones, and Wilks would receive the Annual Lifetime Achievement Awards from the Association for Computational Linguistics in the US.

15. Michael Gavin, "Vector Semantics, William Empson, and the Study of Ambiguity," *Critical Inquiry* 44 (Summer 2018): 653.

Margaret Boden, with her unparalleled knowledge of AI research and cognitive science, shares with us another rare glimpse of Masterman when CLRU first formed in the early 1950s. Like Appiah, albeit twenty years earlier, Boden says that she had adopted Masterman as her mentor and insisted on being taught by this original and somewhat eccentric thinker despite opposition from her director of studies. In the preface to *Mind as Machine: A History of Cognitive Science*, Boden offers a detailed description of CLRU when the group began meeting informally in 1954:

> Masterman's group was doing research on what's now called Natural Language Processing, or NLP. They ranged widely over topics later claimed for AI and cognitive science. These included machine translation, the representation of knowledge for information retrieval, and the nature and process of classification. Although their theory of classification was never described in print as computational "learning", it dealt with issues later so described by AI.
>
> Masterman was one of the first people in the world to attempt machine translation, and she made semantics, not syntax, the driving force. She was deeply influenced by certain aspects of Ludwig Wittgenstein's later philosophy of language. Despite her gender—Wittgenstein was notorious for his mysogyny—she'd been one of his favourite students, to whom he'd dictated the lectures later known as *The Blue Book*.[16]

This last detail caught my eye. Was Masterman Wittgenstein's favorite student? Ray Monk, a biographer of Wittgenstein, appears to confirm the observation. He relates how Masterman and another female student, Alice Ambrose, joined a small group of students Wittgenstein handpicked to attend his course and permitted to take lecture notes. The duplicated set of the notes his students put together was subsequently bound in blue paper covers, hence the title for *The Blue Book*.[17]

Wittgenstein's reputation as a misogynist did not discourage female students from attending his courses, nor did he attempt to exclude them if they lived up to his expectation of honorary males.[18] Although we don't have Masterman's firsthand account of the class dynamic, Ambrose provided a

16. Margaret A. Boden, *Mind as Machine: A History of Cognitive Science*, 2 vols. (New York, 2006), 1:xii–xiii.

17. See Ray Monk, *Ludwig Wittgenstein: The Duty of Genius* (New York, 1990), p. 336.

18. Wittgenstein once said to Elizabeth Anscombe (1919–2001): "Thank God we've got rid of the women!" and was delighted that no (other) female students were in attendance (quoted in ibid., p. 498). Anscombe became one of his closest friends and an editor of his posthumous publications.

description of their experience in her letter to O. K. Bouwsma when the latter sought her story to prepare an article for *The Journal of Philosophy*. I will quote a lengthy excerpt of Ambrose's recollection—not mentioned in Monk's biography for some reason—as it appears in the Bouwsma piece, because it is the sole surviving description of Masterman's interaction with Wittgenstein in that class:

> Wittgenstein was listed in the Cambridge Reporter as giving two courses of lectures in 1933–34, one being called "Philosophy for Mathematicians." To this, as I remember, 30 or 40 people turned up, which distressed him. After three or four weeks of lecturing he turned up at lecture and told the class he couldn't continue to lecture. I remember the occasion and remember how amazed I was that an announced course of lectures could be abandoned in this way. Of the people in that class he chose five of the rest of us to dictate The Blue Book to: H. M. S. Coxeter and R. L. Goodstein, mathematicians, also Francis Skinner (who might have been on a Trinity Grant to do math. though he actually left off doing math. in order to devote himself to Wittgenstein's work), Margaret Masterman Braithwaite and myself. About a month later, I see by a reference to my diary that the five of us had increased to seven, and I know one of them was Mrs. Helen Knight but for the life of me I can't remember the other one. Wittgenstein quarreled with Coxeter because Coxeter quite innocently ran off on a mimeograph the material of the first term's dictation and discussion. So Coxeter didn't continue in the second term. *Mrs. Braithwaite also dropped out during the year in the third term. I've forgotten what the unpleasantness was in her case. She and I took down discussion that he wasn't including in The Blue Book and we called this The Yellow Book. He once flew at her for doing so, but as he was also distressed when something he thought good was not taken down because he wasn't dictating—and she pointed this out to him at the time—* this practice on our part was allowed to continue. I believe I continued with it after she left.[19]

What seems fascinating about Wittgenstein's interaction with his students is not so much his unexpected abandonment of the lecture course as the moments of tension and open confrontation Ambrose has managed to communicate in her letter. Masterman began as one of the teacher's favorite students, but the relationship deteriorated and nearly got out of control. If Masterman dropped out in the third term, was she thrown out like

---

19. Quoted in O. K. Bouwsma, "The Blue Book," *The Journal of Philosophy* 58 (Mar. 1961): 141; my emphasis.

Coxeter? Or did she herself quit? We don't know, and it would be fruitless to push the speculation.

What we do know is that the pedagogical scene in Wittgenstein's classroom—not altogether a successful one—was but a starting point for Masterman, who continued to engage with his philosophical thinking and ended up making an important contribution to ordinary language philosophy after World War II. She is not unacknowledged as the first researcher to introduce Wittgenstein to AI studies, but this recognition says nothing about the originality of her own philosophical work. It is time that we read Masterman closely and carefully to understand how she engages with Wittgenstein's thought in the computing machine and where she departs from him. In a nutshell, we must give her the kind of attention that all formidable philosophers deserve.

Drop the Logos

"What is the meaning of a word?" This opening question in *The Blue Book* is one of many innocuous but difficult issues that Wittgenstein had thrown at Masterman and her fellow students who were tasked with note-taking (*PS*, p. 1). It forces an enduring puzzle upon their attention because the entanglement of word and concept has been a main source of difficulty to philosophers, linguists, translators, and historians. Where does the word end and the concept begin? Is the distinction between word and concept a necessary distinction as Ferdinand de Saussure had insisted? Even if we disregard their distinction in ordinary language use, we still find ourselves wondering what the meaning of a word is and how we are to determine its semantic boundaries.

Having posed his opening question, Wittgenstein goes on to say that when we cannot point to anything in reply to questions like this and feel that we must point to something, we run up against "one of the great sources of philosophical bewilderment: a substantive makes us look for a thing that corresponds to it" (*PS*, p. 1). Saussure, who had developed a theory of semiology in his teachings at the University of Geneva, similarly refuted the correspondence theory of thing and name.[20] He contended that "the linguistic sign unites, not a thing and a name, but a concept and a sound-image."[21] And what is a concept? Is it free from the entanglement attending upon the word in the linguistic sign? Saussure did not pursue such questions—whereas

20. Saussure and Wittgenstein never crossed each other's academic paths in spite of their shared interest. See Roy Harris, *Language, Saussure and Wittgenstein: How to Play Games with Words* (New York, 1990).

21. Ferdinand de Saussure, *Course in General Linguistics*, trans. Wade Baskin, ed. Perry Meisel and Haun Saussy (New York, 2011), p. 66.

Wittgenstein, who may not have been aware of the linguist's work, presses further, adding: "We are unable clearly to circumscribe the concepts we use; not because we don't know their real definition, but because there is no real 'definition' to them" (*PS*, p. 25). Henceforward, starting from his lectures in *The Blue Book*, Wittgenstein begins to develop one of his best-known arguments in the discussion of language games: the meaning of a word (or phrase) is not a mental state or "a mental accompaniment to the expression" but "the use we make of it" (*PS*, p. 65). He argues that there is no such a thing as private language because the meaning of a word happens in the context of language use and will always change depending upon the next context in which the word is used. Wittgenstein would hammer this out in greater detail in *The Brown Book, Philosophical Investigations, On Certainty*, and his other remarks.

Within one year of the publication of *Philosophical Investigations*, Masterman posed her own question "What is a word?" in an article called "Words," in which she calls for a new departure in ordinary language philosophy by interrogating the identity of *word* itself.[22] The deliberate shift from "meaning" to "word" in the wake of Wittgenstein's question allows her to reopen one of the seemingly indisputable linguistic facts to philosophical inquiry. Her question does not require a new definition of *word* any more than does Wittgenstein's earlier question requires a better definition of the "meaning of a word." It centers, instead, on the undecidability of the identity of *word* in language use.

Take one of the examples she introduces into the discussion. How do we know that *ward* ("a person or minor under protection"), *ward* ("room in a hospital"), and *ward* ("to parry" in fencing), and many other usages in English are one and the same word with different shades of meaning rather than, say, different words that happen to share the material sign *WARD* with respect to phonemes (homophones) and spelling? The *Oxford English Dictionary* (*OED*) has simplified the matter by classifying the multiple uses of *ward*—as verb, noun, adjective suffix, adverb suffix, proper name—under one entry, essentially treating them as a single word under the rule of polysemy.[23] This solution is convenient but proves incapable of resolving the undecidability of *word* as a philosophical conundrum.

That conundrum is by no means an idle issue. Masterman's research group at CLRU learned it the hard way within a couple of years when they embarked on the computational research on machine translation and information retrieval. The undecidability of *word* in either determination—a single word with multiple senses or multiple words unified by a single sign— became an endless source of frustration and challenge for them, trumping all

22. Masterman, "Words," in *Language, Cohesion and Form*, p. 21; hereafter abbreviated "W."
23. *Oxford English Dictionary*, s.v. "ward."

other difficulties. The *OED* turned out to be the least helpful template when the researchers ran into technical difficulties while mining data from pedestrian language use or when they found themselves overwhelmed by the ubiquity of word-concept entanglement in the machine. I take up the technical issues and solutions in the last section.

If the conundrum of word and concept is more of a philosophical problem than it is linguistic, lexical, or technical, should we look to philosophers for a good answer? Masterman, however, remains skeptical of the postwar generation of philosophers, surmising that these men would most likely give one of the three answers. The first is that everybody knows what a word is; the second is that nobody knows what a word is; and the third or the typical reaction from logicians and analytical philosophers is that "it doesn't matter anyway what a word is, since the statement is what matters, not the word" ("W," p. 31). Masterman tries to show why this argument is deeply flawed. She points out that "the logical importance of finding out what we think about words—and the logical importance of examining and distinguishing usages in context—turns out to lie in the fact that our conception of a 'statement' (and, a fortiori, of logically possible forms of connection between 'statements') is fundamentally affected by our logical conception of a 'word'" ("W," p. 38). Herein lies the philosophical impasse: The logical conception of word—in contrast to the common dictionary definition of *word* with its tautologies—is responsible not only for the possibility of a statement but also for the word-concept entanglement that the logical statement is expected to clarify or undo.

Therefore, the critique of the logical conception of word constitutes the first step toward a post-Wittgensteinian philosophy of language. "In entering this new universe of discourse," Masterman proposes, "the logician has to prepare himself for two successive and contrary shocks." The first is the shock of finding a "very great deal of indeterminacy everywhere" (Wittgensteinian argument), and, after one has faced the first shock, the second is the shock of finding "what unforeseen new vistas open out, and what a lot can be done," which foreshadows the post-Wittgensteinian machine she would be building at CLRU. In her judgment, philosophers of ordinary language who take an interest in formal logic have made "the mistake of putting their positive and negative results much too much at the service of the old logical approach, in order, apparently, to try and sophisticate it." This attempt is fruitless. What we need is "a new sensitiveness to ordinary language" and "a fundamentally new approach to the problem of what logic is" ("W," pp. 36 n 3, 36–37 n 3, 37 n 3).

Is there anything new about Masterman's attack on the old logical approach? One of her objectives is to overcome what Jacques Derrida later termed *Western logocentrism*. In that sense, her work may be said to anticipate

the French philosopher's critique of metaphysics, but this is not the argument I am trying to make here. Rather, I want to stress the essential differences: For Masterman, to overcome Western logocentrism means opening up the ideographic imagination beyond what is possible by *the measure of alphabetical writing*. This is important, as it follows that the scientist's and philosopher's reliance on conceptual categories derived from alphabetical writing in their commitment to logical precision and systematization as well as their deconstruction must likewise be subjected to post-Wittgensteinian critique. This philosophical ambition is amply reflected not only in Masterman's published essays from the early 1950s onward but is implemented, methodically and painstakingly, in her innovations of MT and IR technologies in the subsequent decades. Her engagement with philosophy and, in particular, with Wittgenstein's later philosophy remains inseparable from the work of CLRU on computer algorithms. For this reason, we are fully justified to describe Masterman's work as *doing philosophy in the machine,* literally speaking. What she did was turn the cognitive limitations of the computer—that is, the challenges involved in the programming of the MT machine to distinguish amongst *ward*, *ward*, and *ward* as different words or as different senses of one word—into a distinct advantage to achieve greater philosophical clarity about the entanglement of word and concept in human languages.

To that end, Masterman seizes upon the logical conception of word as her main target in the critique of Western logocentrism. To push the critique beyond what is possible by *the measure of alphabetical writing,* she takes the bold, preliminary step of opening up the ideographic imagination for future philosophy. This is already present in an earlier albeit shorter piece she had written and presented at the International Congress of Philosophy in Brussels on 20–26 August 1953 where she began to elaborate a "pictorial principle" that opposed the ideographic conception of language to logocentrism. In her paper "The Pictorial Principle in Language," Masterman accuses formal logicians—known as analytical philosophers in the US—of reducing "thought" to the manipulation of logical units called statements. Their procedure runs like this: Statements are at their most statement like when they are deducible from other statements; and statements are at their most elegant when they are systematized. She goes on to observe:

Now, though, when we see the advantages which the successive establishment of precision, deductive connection and systematisation have produced in science after science, we feel that, in a sense, the logicians are right in their estimate of "what thinking really is", the fact remains that when we ourselves think most deeply, *we nearly always*

throw the whole logical machinery over. At such times we "doodle", we
compare, we "match"; we write down isolated words, we draw pictures on
the edge of the paper, we make models.[24]

In the italicized quote, the author cites seemingly random acts as examples of pictorial thinking: doodling, comparing, matching (patterns), scribbling, drawing pictures, and so on. These acts are not as random as they first appear and would prove integral and indispensable to the thesaurus method Masterman would develop for machine translation at CLRU. Almost by necessity, she and her group resorted to scribbling, doodling, matching, and hand-drawn diagrams and pictures because the computer of the mid-1950s was rudimentary, and much of their early work had to be done first by hand and then performed on Hollerith punched-card machines. The first digital computer did not arrive in the office of CLRU until 1963 or 1964, and it was a primitive ICT1202 with only 4K storage and no backup.[25]

More fundamental than the technological handicap is the philosophical question of what constitutes pictorial thinking. Masterman dismisses the notion that pictorial thinking has anything to do with the primitive vestigial habit or that it is about pictorial forms of representation. She relates it, instead, to Wittgenstein's idea of the logic of representation in his earlier study of language in *Tractatus Logico-Philosophicus*. To carry that work forward, Masterman proposes a series of methodological interventions:

The first is that we must develop exact analytic procedures, in order to discover what Wittgenstein, speaking more literally than he knew, called "the logic of representation". The second is that the proper field in which to attack this problem is not that of the psychological study of images or the ethnological study of sign-stimuli, nor that of the philosophy or psychology of visual artistic creation, but that provided by the well-documented existence of an actual language founded upon what I shall call "the picture principle"; that is, by investigating the logical forms in Classical Chinese. ["PP," pp. 139–40]

This is an unorthodox proposition, one that few have contemplated or thought worthwhile except for a handful of mathematicians and philosophers who urged a similar investigation, including Ernst Mach, Alfred Whitehead,

---

24. Masterman (Braithwaite), "The Pictorial Principle in Language," *Proceedings of the Xlth International Congress of Philosophy* 14 (1953): 139; my emphasis; hereafter abbreviated "PP."

25. See Masterman, "Man-Aided Computer Translation from English into French Using an On-line System to Manipulate a Bi-lingual Conceptual Dictionary, or Thesaurus," in *Proceedings of the 1967 Conference on Computational Linguistics*, 23–25 Aug. 1967, p. 15. See also Wilks, "Margaret Masterman," p. 282.

and Friedrich Waismann, whose names appear in passing.[26] Why classical Chinese? It is not that Masterman thinks that the ancient writing system is a collection of pictures or pictographs—an unfortunate caricature disseminated by early Christian missionaries—but that the ideographic writing operates on combinatory logic, not propositional logic. The difference between combinatory logic and propositional logic carries tremendous importance for her. To investigate the logical forms in classical Chinese is to look for the rules of combination of ideographic clues or visual hints, and this is conceivable only when one ceases to think of ideographs in classical Chinese as pictorial representations of objects or icons by resemblance (see "PP," p. 140). Masterman's method is motivated by her discovery of the ideographic asymmetric sequence in classical Chinese where "logical connections are made by creating and by combining combinators,—and nearly all these combinators, in their primary function as elements, have demonstrative (i.e. indexical) significance" ("PP," p. 142).

Take a typical Chinese phrase with a string of adjectives followed by a noun. She determines that the string of adjectives is not arranged haphazardly, for the rule that governs it—logical, not grammatical—consists of an ordering of conceptual elements going from the more abstract to the more concrete; from the more general to the more particular; from the more universal to the more limited. This combinatory logic takes the form of an ideographic asymmetric sequence such as, $(a(b(c)))$, which she spells out as follows: "the idea of c, limited, or qualified by the idea of b, the whole, c-qualified-by-b, being further qualified by a," and the sequence can be extended on and on. In mathematical terms, she terms it Ascriptive Combinator or A-Combinator and concludes that "it is ordering elements in this way which I call the pictorial principle in language; and it is an extremely fundamental and pervasive principle in all languages" ("PP," p. 143).

That is to say, the combinatory logic of ideographic writing is not confined to classical Chinese that happens to embody it. The same logic can be extended to a new understanding of English and other languages until the ideographic principle is shown to operate in all languages. This universal principle, more logically fundamental than the measure of alphabetical writing, is poised to challenge the metaphysical foundation of the philosophy of language. To clinch the point, Masterman set out to work on a substantial philosophical treatise with the title "Metaphysical and Ideographic Language" to be published in

26. Although Masterman made an effort to study classical Chinese, she sought help mainly from linguist M. A. K. Halliday (Michael Halliday) and followed the work of sinologist Gustav Haloun and Y. R. Chao. On Halliday's role in CLRU, see Wilks, "Margaret Masterman," p. 280.

British Philosophy in the Mid-Century (1957).[27] At the same time, she began to subject the combinatory logic of ideographic writing to the operational test on the computer.

Before moving on to the computable thesaurus, we need to probe her ideographic principle further. My discussion began with Masterman's opening question in "Words": What is a word? Following upon her critique of logocentrism in "The Pictorial Principle in Language," a slightly different question emerges: In what ways does her ideographic principle help illuminate the philosophical conundrum of word-concept entanglement in language? For instance, when the MT researcher is confronted by the undecidability of the identity of *ward, ward, ward,* can she rely on the ideographic principle and its combinatory logic to determine whether she is looking at a single word with different meanings (polysemy) or whether these are multiple words that are somehow unified by a single written sign *WARD*? In anticipation of Masterman's surprising answer, we need to place the question on slightly different footing to ascertain its relevance to the central concerns of ordinary language philosophy raised first by Wittgenstein.

Word, Pattern, and Ideograph

Consider Wittgenstein's discussion of the numeral in *The Brown Book.* This is where he introduces a distinction between word and pattern before casting doubt on it. Wittgenstein begins:

> We may say that words and patterns have different kinds of functions. When we make use of a pattern we compare something with it, e.g., a chair with the picture of a chair. We did not compare a slab with the word "slab". In introducing the distinction, 'word/pattern', the idea was not to set up a final logical duality. We have only singled out two characteristic kinds of instruments from the variety of instruments in our language. We shall call "one", "two", "three", etc., words. If instead of these signs we used "–", "– –", "– – –", "– – – –", we might call these patterns. Suppose in a language the numerals were "one," "one one," "one one one", etc., *should we call "one" a word or a pattern?* The same element may in one place be used as word and in another as pattern. [*P*, p. 84; my emphasis]

A marvelous thought experiment, one that effectively challenges our preconceived ideas about language. Still, Wittgenstein's hypothetical numerals,

27. This seventy-eight-page-long treatise deserves separate treatment that I will attempt in a forthcoming book-length study. See Masterman, "Metaphysical and Ideographic Language," in *British Philosophy in the Mid-Century: A Cambridge Symposium*, ed. C. A. Mace (London, 1966), pp. 283–357.

"one," "one one," "one one one" will give us pause in spite of our willingness to accept the argument that *one* may in one place be used as word and in another as pattern. The difficulty is how we can wrap our mind around the possibility that a word written in alphabetical writing like *one*—as opposed to individual letters—could ever conceptually become a pattern or nonword, although the reverse happens more plausibly as in the case of rebus. In my view, it is one thing to argue that the meaning of a word is its use depending on the circumstance and quite another to suggest that a word can become a nonword (if we suspend nonsense spelling and typographical experiment for the time being), assuming that the distinction of word and pattern will hold. Wittgenstein's hypothetical numerals raise the possibility of the ideographic determination of alphabetical writing, a scenario more radical than he would have imagined.

To elaborate, I offer a couple of additional observations. First, beyond what Wittgenstein says about the ambiguity of word and pattern, there is the undecidability of the one and the many that troubles the second and third of his hypothetical numerals. For instance, how can we determine whether "one one" should be taken as two words/one pattern or two words/two patterns expressing a single concept like that of Hindu-Arabic numeral 2? Concomitantly, are we supposed to take "one one one" as three words, one pattern, or three patterns, all denoting a single concept as does numeral 3? It seems that the problem of the one and the many refuses to go away. This predicament bears upon Masterman's case of *ward, ward, ward* and the undecidability of their word/pattern makeup.

That brings me to my second point. In his thought experiment, Wittgenstein proposes two types of hypothetical numerals: "one," "one one," "one one one" versus "–," "– –," "– – –," and so on, and he calls the horizonal bars "patterns." The pattern is construed as a nonword and negatively determines the identity of word. Even if we preclude the thought of "final logical duality" as Wittgenstein rightly cautions us, we are still forced to determine the identity of pattern by adhering to the metaphysical distinction of word and nonword. Once we set on this path, is there a limit to which we may extend the negative reciprocal determination of word and pattern? Take Roman numerals I, II, III where the alphabetical letter *I* is borrowed to form numerals. The numerals cannot themselves be words in order for them to be simultaneously associated with English words *one, two, three*, French words *un, deux, trois*, or Mandarin *yi, er, san*, or any other linguistic system. Following Wittgenstein's argument, are we to call these numerals "patterns" formed with alphabetical letters, not unlike his hypothetical "one," "one one," "one one one"? Alternatively, we might rewrite the same numbers as 一, 二, 三, or 〡, 〢, 〣 by adopting one or another of established systems of Chinese numerals, the latter

set being ancient and obsolete. These numerals are likewise associated with multiple linguistic systems across East Asia, such as Mandarin, Cantonese, Min languages, Japanese, Korean, and others where the numerals cannot simultaneously be words. Are these patterns instead? Compared with Roman numerals, the Chinese numerals bear closer visual resemblance to Wittgenstein's patterns "–," "– –," "– – –." Can we conclude that they are more pattern-like than I, II, III? On further reflection, the comparison between the Roman numerals and the Hindu-Arabic ones (1, 2, 3) gives the impression that the former is visually more pattern-like than the latter. What is a pattern anyways? The impasse we are grappling with forces us to consider whether visual criteria should be used—as seemingly implied in Wittgenstein's example—to adjudicate the metaphysical distinction of word and nonword.

Precisely to address such metaphysical predicaments, Masterman makes a startling move in "Words" to initiate a series of inquiries into the ideographic determination of alphabetical writing. She does so by developing an "exact analytical procedure" for philosophical analysis that would mark a bold departure from Western metaphysics and the beginning of post-Wittgensteinian philosophy. Central to her procedure is what she borrows from classical Chinese called the *zi* 字 (tzu⁴ in her spelling; pronounced *ji* by the Japanese as in *emoji*).[28] For Masterman, the *zi* is what makes the general and abstract category of the written sign possible, for not only does the *zi* override the Wittgensteinian distinction of word and pattern, but it also renders the distinction of word and nonword superfluous.

To demonstrate her new method, Masterman proposes a language game in which the four-letter sequence *ward* is made to behave like a *zi*. The *zi ward*—being a logical unit rather than a syntactic unit—momentarily suspends the metaphysical question as to whether *ward* should be taken as a single word with many meanings—that is, polysemy—or as several words that happen to share the same letter sequence. Instead, she focuses on the total spread of the *zi* to map out an indeterminate sequence that begins with *ward* in isolation and becomes cumulatively determinate as complication grows in context. The semantic spread of the *zi* does not coincide with what we call *word* or *words* in English nor does it depend on grammatical units. With a sufficiently long unit of discourse, some rules for distinguishing one

---

28. The neologism *emoji* transcribes three *kanji* characters in Japanese or Chinese: 絵 ("picture") 文字 ("written character"), the last 字 being pronounced *zi* in Mandarin and *ji* in Japanese. Masterman relied on the Wade-Giles romanization scheme to write the same character as tzu⁴. The collection from which I cite this work has introduced a typo *tzu4* that I correct as *tzu*⁴ here on the basis of Masterman's original journal article. The superscript 4 in tzu⁴ denotes the falling tone in Mandarin. I adopt the standard Pinyin romanization *zi* when referring to her work while keeping her original spelling tzu⁴ in quoted texts.

usage from another under the ideographic sign of the *zi* can be devised; which is to say, one can make semantic distinctions without getting caught up in the metaphysical entanglement of word and concept.

Masterman's language game works thus: Instead of looking up the polysemic word *ward* in a dictionary, we are asked to consider the behavior of the *zi ward* in the following sequence: "'I ward thee a *ward*, Ward. WARD!- *Ward WARD*ward, Ward! Ward not to be warded, Acreward, in *ward*, I ward thee, *WARD!* WARD! WARD! *I ward* thee, *WA-ARD!*'" ("W," p. 56). This thought experiment borders on Joycean exuberance in polysemy, but what she is really doing is repeating the *zi ward* eighteen times and use punctuation, typography, and other ideographic marks to provide each occurrence with a context and a specific sense. The quoted discourse, being sufficiently long and sufficiently complex, gives the speaker of English a rough sense of how the speaker warns someone named Ward to defend himself by parrying the other man's blow with the ward of his sword and further warns him that if he doesn't do this, he (Ward) will find himself carried off to Acre and there cast into a prison cell, and so on. Whichever direction in which one may choose to interpret the meaning of each use of *ward*, there emerges in her language game almost a vectorial principle governing the cumulative determination of indeterminate units in ordinary language use. Each repetition is a repetition with difference. What does the repetition with difference entail? Masterman replies: "The fact that formal logical games are played with indeterminate units does not prevent specific rules from being devised to govern the methods of combination of such units" such that "the games can be so played that, as the sequence of units extends, the spread of the sequence becomes more and more determinate; as complication increases, indeterminacy grows less" ("W," p. 35). This appears to support the Wittgensteinian argument that the meanings of a word are never independent of the contexts in which it is used.

One could object that there is nothing new about the determination of a word's meaning by context. Should we even bother about the ideographic *zi* when we may rely on our intuition to comprehend all of this? The answer lies not in our intuition but in the formal mechanism—the logical mechanism or algorithms—that can predict what we comprehend through our intuition and, therefore, lead to a robust theory of how particular meanings emerge from linguistic context. What Masterman has done is string together eighteen different occurrences and uses of the *zi ward* and punctuate the string by inserting stress marks, capital letters, a suffix (*-ed*), as well as a few connecting phrases. More importantly, she adopts what is called logical "bracketing" to join verbal units in Ward*ward* or bring several *wards* together, with a comma, an exclamation mark, or a period or to bind three *wards* together by

increasing the stress ("W," p. 36). In lieu of syntactical units, Masterman speaks of small words or other devices that are used as operators to define the interrelations of these brackets. To render the interrelations of the brackets fully legible, she transcribes the eighteen-*ward* language game in a coded asymmetric sequence thus:

[[[/ ((({W}(W))W)F) / ((W)E) / (((W(WW)W))E) ///]
(((W((NW)P)) ((AW)IW)) ((({W})W)E)) ///]
/ ((W)E) / ((W)W) / ((({W})W)E) ///]
["W," p. 36]

Here she introduces a distinction between the conventions for the ideographic *zi* as formulated in table 1 and the conventions for what she calls "phrases and phrasing" (to distinguish her logical unit from the syntactic unit) in table 2 (fig. 1). The tables explain the ideographic symbols Masterman uses to elucidate the interrelations of the brackets in the logical formulation and its eighteen reiterations of the *zi ward*. "In so far as we have adopted as logically fundamental that analytic technique that consists in comparing, in context, slightly differing usages and uses of the same word," writes she, "just so far *we have already committed ourselves, though we may not know it, to a technique that works in terms of tzu⁴, not in terms of words*." ("W," p. 31; my emphasis).

Clearly, the ideographic asymmetric sequence of *ward* is much more than a thought experiment with polysemy, because the "technique" she introduces contains the seeds of radical philosophical movement between alphabetical writing and ideography. This is what I meant when I said that the overcoming of Western logocentrism did not mean the same thing for her as it did for Derrida, as I am fully convinced that Masterman is the first modern philosopher to push the critique of Western metaphysics beyond what is possible by the measure of alphabetical writing, and, unlike deconstruction, her translingual philosophical innovation refuses to stay within the bounds of self-critique.

Driven by a shared ideographic script associated with multiple languages across East Asia, Masterman determines that the semantic spread of the discrete *zi* vastly exceeds any known relations of word and concept in alphabetical writing, because the *zi* is the site where multiple languages (and numerals) converge and crisscross or "interlingual both in space and time, that is denoted by a single ideographic 'character', that is, by a visual shape, or graph" ("W," p. 28). To imagine a graph in English that approximates the semantic spread of the *zi* is to ask a word like *ward* to stop being itself and become manifold under the sign of its own written form as a generalized *zi*, more abstract and universal than logos itself. This is the philosophical task she seeks to accomplish. By extending the ideographic principle of the *zi* to other languages

Table 1. *Conventions for tzu4*

tzu4	*Abbreviation*
Ward	W
Acre	A
Finality (assertion)	F
Exclamation	E
Inclusion	I
Passivity	P
Nonning (cancellation)	N

Table 2. *Conventions for the phrases and phrasing*

Phrases/phrasing	*Abbreviation*
'I … thee'	{ … }
'a … '	(…)
' … punctuation sign'	(…).)
' … , … '	(…) …)
'_.===,_'	/ / / ///
' … and … '	[[ ]]

Note: When two tzu4 are immediately juxtaposed, they shall be interpreted as bracketed thus: ((.) .)

FIGURE 1.   Tables from Masterman, "Words," p. 37.

and generalizing it in the machine, Masterman's philosophy succeeds in pushing beyond Wittgenstein's speculation on word, pattern, and picture.

What does this move signify? I would say that it signifies a fundamental repudiation of the unity of word and concept or that of icon and concept. Of course, neither Masterman nor Wittgenstein was the first to critique these unities. Gottlob Frege, among others, had rejected any attempt to equate

words with concepts. In his formulation of logical statement in language, a concept fulfills the grammatical function of the predicate more or less commensurate with its mathematical function, and it shows up in the (unsaturated) position of the predicate but never in the position of a grammatical object.[29] The distinction he maintains between object and concept is helpful for the purpose of symbolic logic, but the logical statement is precisely where Masterman begins her departure from the formalism of logicians who "tend to regard 'thought' as the manipulation of logical units" and whose symbolic manipulation is so strictly circumscribed that it leaves the actual language data or usage unexamined ("PP," p. 139).

It is a mistake to conclude, nonetheless, that Masterman's critique of logical statement is a rejection of logic or logical reasoning. Far from it, she expresses a keen interest in developing an ideographic principle of language whose combinatory logic prioritizes the logical unit of language over the syntactic unit (see "PP," pp. 140–41). She further claims: "It can be shown that if we once adopt as our logical unit the logical conception of a tzu[4], for instance, rather than that of a word, we take a step that has fundamental logical effects" ("W," p. 31). With the *zi* taking precedence over word, her methodological foray leads to the next step of approaching English writing as an ideographic system, not unlike what Claude Shannon did when he introduced "space" as the twenty-seventh letter of Printed English in information theory.[30]

In "Words," Masterman finds herself engaging extensively with the work of Yuen Ren Chao, an esteemed linguist based at the University of California, Berkeley and onetime president of the Linguistic Society of America.[31] Chao was the first to identify the conceptual gap between the *zi* and the "word," one that poses "the identity-of-unit problem: What constitutes one and the same word?"[32] He had long grappled with the widespread confusion surrounding the semantic unit of the Chinese language whenever the category of word is applied in linguistic and lexicological studies. Chao decided to

---

29. Frege did most of this work in three influential essays "Function and Concept," "On Sense and Reference," and "On Concept and Object"; see Gottlob Frege, *Translations from the Philosophical Writings of Gottlob Frege*, trans. P. T. Geach et al., ed. Geach and Max Black (Totowa, N.J., 1980).

30. For my discussion of Shannon's invention of ideographic English for information systems, see Lydia H. Liu, "The Invention of Printed English," *The Freudian Robot: Digital Media and the Future of the Unconscious* (Chicago, 2010), pp. 39–98.

31. Chao had taught at Tsinghua University in Beijing. Among his better-known maverick feats was the first Chinese translation of Lewis Carroll's *Alice in Wonderland* (1922) and his support of I. A. Richards's promotion of BASIC English in China before WWII.

32. Yuen Ren Chao, "The Logical Structure of Chinese Words," *Language* 22 (Jan.–Mar. 1946): 4.

adopt the *zi*—the written sign, not "the syntactical word"—as a unit of analysis in his 1946 essay titled "The Logical Structure of Chinese Words" where he points out that the *zi* articulates the logical structure of Chinese "words" or "free form."[33] The fraught relationship between the written character and its multiple linguistic counterparts is borne out by the phenomenon of mutually incomprehensible pronunciations of the same character across many languages in East Asia.[34] This diversity and overwhelming tension is exacerbated by millennia-long morphological changes and linguistic borrowings that result in increased homophones and synonyms in a writing system shared by diverse languages. Prior to the introduction of English grammar and grammatical texts from Europe, the *zi* had dominated all conceptions of language in China and East Asia to the extent there was no word for *word* nor was it deemed necessary to invent one until the early twentieth century.[35]

Chao's insight led Masterman to question the grammatical word as a linguistic fact. "Speaking grammatically," she writes, "it is an understatement to say that it is not clear what a grammatical word is. What is clear is that Chao's distinction between a 'free form' and a tzu^4 [*zi*] is logically fundamental" ("W," p. 31). In lieu of logos, a written character corresponds to as many semantic and/or phonetic units (some of which coincide with what we call *words*) as there are languages and dialects. The *zi*, therefore, epitomizes the problem of the one and the many between (one) writing and (many) languages that constantly calls for contextual determination of meanings.[36]

How unique is the situation of the *zi*? Is it not true that all communicative acts require contextual determination, regardless of languages and writing systems? Yes and no. It depends on what we mean by context. In her reading of Chao, Masterman states that "we ourselves have tzu^4 [*zi*], or general root words, in English; *ward* is one. With a little trouble, we could compile a glossary of the most fundamental English tzu^4, and we could then invent ideographic signs for them" ("W," p. 28). The bold translingual move she proposes seeks to overcome the philosophical gap between the *zi* and the "general root word" in English. "It is extremely difficult to exemplify this process in English," she says, "since we have become accustomed to think

33. Ibid.

34. See ibid., p. 10. The same holds for the writing systems in Japan, Korea, Vietnam (before 1911), and other systems where the ideographic script (*kanji* in Japanese) has been adopted or partially adopted.

35. The first Chinese grammarian Ma Jianzhong translated the English term *word* as *zi* in 1898, which caused confusion and contentious debates among Chinese linguists. Subsequently, *word* was retranslated as *ci* (a super-sign) to help negotiate the gap between the *zi* and the *ci* (word). For detailed discussion, see Liu, "The Sovereign Subject of Grammar," *The Clash of Empires: The Invention of China in Modern World Making* (Cambridge, Mass., 2004), pp. 181–209.

36. See Chao, "The Logical Structure of Chinese Words," pp. 10–11.

of English as built up out of words, not of tzu[4]" ("W," p. 35). And what does it mean to imagine English words as the *zi,* using the ideographic technique?

It means that, in place of metaphysical distinctions, the *zi* enables semantic phrasings, clusters, and classes on the basis of combinatory logic. As a generalized, abstract category of inscription, the *zi* is capable of subsuming different classes of signs, such as the *shuzi* (numerical *zi*) and the *wenzi* (textual *zi*). The *shuzi* encompasses a multiplicity of indigenous and imported numerical systems such as 一, 二, 三; ｜,｜｜,｜｜｜, or 1, 2, 3, irrespective of patterns, words, or symbols. The *wenzi,* on the other hand, consists of characters used in composing texts, although the written character—called the syllabic morpheme by linguists—is not a word unit by any measure. Thus, Masterman translates the abstract, universal principle into a set of sophisticated computational techniques to guide the work of the CLRU where she and her team developed a computable thesaurus and an interlingua for machine translation.[37] She declares: "On all sides, fascinating logical vistas open out before me. I stand, like Alice, above a chessboard which covers the whole world" ("W," p. 37).

Masterman's chessboard is no ordinary chessboard. Overlaid with allusions and extended metaphors, it is associated, on the one hand, with Wittgenstein's analogy of words as chess pieces and, on the other, with the computer programs that she and her team would construct at CLRU. Within a decade or so, her machine would leave the analogy of words and chess pieces behind to develop automated information retrieval, machine translation, and other computational programs to be associated with machine learning.

Computable Thesaurus

Imagine the Wittgensteinian language game being played on a corpus as enormous as *Roget's Thesaurus* or the entirety of English vocabulary, and you get the idea of what it takes for Masterman to construct a machine capable of handling the fraught relationship of word and concept in human languages. Will her post-Wittgensteinian machine live up to the ambiguity and multiplicity of word meanings, cross-references, analogies, and metaphors in language uses?

In what follows, I undertake a close analysis of Masterman's elaboration of the technique that "works in terms of tzu[4], not in terms of words" as she attempts to achieve her philosophical goals in the machine. Up till now, I

---

37. Research at CLRU took off officially when the group received major grants from the National Science Foundation totaling 101,250 dollars between 29 March 1957 and 6 May 1960; see Automatic Language Processing Advisory Committee (ALPAC), *Languages and Machines: Computers In translation and Linguistics* (Washington, D.C., 1966), p. 107.

have focused on her contributions to the philosophy of language and have not been able to take up her specific innovations in MT and IR. The truth is that these two aspects of her work—philosophical and computational—are inextricably interwoven and should be examined together. One cannot speak about her innovation of mechanical thesaurus and semantic networks without taking into account her critique of logical calculus associated with Bertrand Russell and mainstream analytical philosophy. Although the story of her pioneering role in NLP is well known to AI practitioners, the news of her philosophical breakthrough has not arrived despite the fact that her avowed goal was to achieve greater philosophical clarity on the question of language, a promise she amply fulfilled by doing philosophy in the machine. What does this *doing* entail? We will have a glimpse of it by examining the mechanical (computable) thesaurus that she designed with her team at CLRU.[38] The thesaurus method—worked out philosophically through her explicit engagement with Wittgenstein's later work—anticipated word-sense disambiguation (WSD) and vector-space semantics in the AI research of our time.

Commenting on the computational study of "word-sense disambiguation," Gavin has credited its recent success to vector-space semantics, a popular method in NLP that works by finding clusters in the semantic space of a word and measuring how closely any individual use of the word sits near each cluster. He shows that vector-space semantics is "a theory of ambiguity, pushing strongly against the impulse to draw clear boundaries that isolate words into discrete concepts."[39] Tracing the evolution of this method, Gavin shows that Masterman and her team developed the first computer-based thesaurus drawn from *Roget's* when they modeled word meanings not in terms of word-concept correspondence but on distributed processes across a corpus in lattice-shaped network in the 1950s.[40] I agree with this assessment but must add that Masterman's computable thesaurus, in particular, is the undisputed forerunner of word-sense disambiguation, even though some of the breakthroughs had to wait till the 1990s when latecomers such as David Yarowsky tapped into *Roget's Thesaurus* again to develop sophisticated algorithms for unsupervised machine learning on faster and more powerful computers.

The system of classification in *Roget's Thesaurus* has some distinct advantages for the computing machine over the *OED* that privileges etymologies

38. CLRU relied on mechanical means of drawing up paper lists and using punched card apparatus to process the lists. Their mechanical thesaurus consists of programs or semantic algorithms that are computable and, therefore, may be called a computable thesaurus *avant la lettre*.

39. Gavin, "Vector Semantics," p. 659. Known as distribution hypothesis in NLP, it helps determine morphological segmentation.

40. See ibid., p. 653.

and definitions. The thesaurus identifies patterns and linkages among words and phrases and employs a numerical system to assign the proximate location of each use in relation to other locations.[41] Wilks recalls Masterman stating that "thesauri like Roget were not just fallible human constructs but real resources with some mathematical structure that was also a *guide to the structures which humans process language.* She would often refer to '*Roget's unconscious*' by which she meant that the *patterns of cross-references, from word to word across the thesaurus*, had generalisations and patterns underlying them."[42] AI scientists today would be inclined to translate her term "Roget's unconscious" into neural networks.

If the *OED* is philological and linguistic, the thesaurus is philosophical and mathematical where the meaning of a word is not defined by other words so much as determined by what surrounds its use or by the location where the word occurs (its context) with a certain pattern or regularity. The MT program modeled on the thesaurus must be able to "face and not evade the problem of the indefinite extensibility of word meaning," observes Masterman.[43] It also means embracing and not eliding the ubiquity of metaphors, analogy, poetry, and the language games with which Wittgenstein tried to grasp the workings of ordinary language. For instance, Masterman had experimented with a computer program written in the TRAC language to produce machine-generated Japanese haiku poetry in 1968 (see "S").[44] The novelty of a thesaurus as a model lies in its potential for analyzing the indefinite extensibility of word meaning as a philosophical problem, and "it ties up my translation model not to philosophy in general, but to a particular kind of contemporary philosophy, namely, linguistic philosophy, the 'philosophy of ordinary language'" ("T," p. 187).

Wittgenstein is known for his use of the duck/rabbit drawing to illustrate how two mutually exclusive meanings arise out of a single icon. A related but more powerful illustration is his analysis of the multiple aspects of the icon △ that can be read "as a triangular hole, as a solid, as a geometrical drawing; as standing on its base, as hanging from its apex; as a mountain, as a wedge, as an arrow or pointer, as an overturned object, which is meant . . . to stand on the shorter side of the right angle, as a half parallelogram, and as various other

---

41. See Masterman, "What Is a Thesaurus?" in *Language, Cohesion and Form*, pp. 134–38.

42. Wilks, "Editor's Introduction," in *Language, Cohesion and Form*, p. 7; my emphasis.

43. Masterman, "Translation," in *Language, Cohesion and Form*, p. 187; hereafter abbreviated "T."

44. For her machine-generated haiku, see Robin McKinnon Wood and Masterman, "Computer Poetry from CLRU," in *Cybernetic Serendipity: The Computer and the Arts*, ed. Jasia Reichardt (New York, 1968), p. 55.

things."[45] Wittgenstein suggests that what we call a triangle in language or what we observe in the icon itself does not always correspond to a single concept. The visual image △ refers to a set of family resemblances that may overlap but do not necessarily coincide with one another, and the same thing happens to the linguistic abstraction *triangle*. If this insight appears to resonate with Masterman's analysis of the word *ward*, it is because she never ceases to engage with Wittgenstein. The diagram represents one of her attempts to formalize Wittgenstein's discussion of the icon △ in her philosophy machine (fig. 2).

In dialogue with Wittgenstein, Masterman's diagram enacts the shape of which it speaks by placing the visual sign △ at the apex to map out a triangular network of family resemblances amongst the multiple uses (meanings) of smaller icons that run parallel along the base. The big icon △ at the apex of the diagram can be taken as a visual equivalent of the "head" in the computable thesaurus.[46] The diagram forms an icon—which Masterman terms a "fan"—of the iconic relations in a network of family resemblances it seeks to demonstrate.[47]

To visualize Wittgenstein's philosophical argument about a particular icon is not the same thing as to demonstrate how a network of family resemblances—iconic or otherwise—operates on the basis of the ideographic/mathematical principle. W. J. T. Mitchell has interpreted the Peircean icon in similarly ideographic terms: "At the heart of logic and mathematics, then, the iconic relations of identity and equivalence, similitude and difference, are lurking."[48] Consistently visual and spatial, Masterman's semantic patterns derive from a finite set of semantic primitives or "sticking figures" she had adapted from I. A. Richards and Molly Gibson's *Language Through Pictures* book series (see "T," pp. 171–73).[49] She presents them as iconic figures known as fans, semantic shells, lattices, and semantic squares in the computable thesaurus. "In order to get semantic patterns on to a machine," says she, "we have created in CLRU a unit of semantic pattern called a template" and so on, a project that others like Sowa and Yarovsky would pursue and advance decades later ("S," p. 261). As she explores the networks of overlapping and criss-crossing word meanings or semantic

45. Wittgenstein, "Philosophy of Psychology—A Fragment," *Philosophical Investigations*, p. 210e.

46. For her comparison of the thesaurus approach to language uses and Wittgenstein's gestalt-theory of concept, see Masterman, "Fans and Heads," in *Language, Cohesion and Form*, pp. 39–53.

47. See ibid.

48. W. J. T. Mitchell, *Image Science: Iconology, Visual Culture, and Media Aesthetics* (Chicago, 2015), p. 29. There is evidence that Masterman read Peirce, but she rarely refers to semiotics.

49. The series was created in 1952–1958 to promote the teaching of Basic English. Their technique consists in portraying basic situations in real life similar to that used in comic strips.
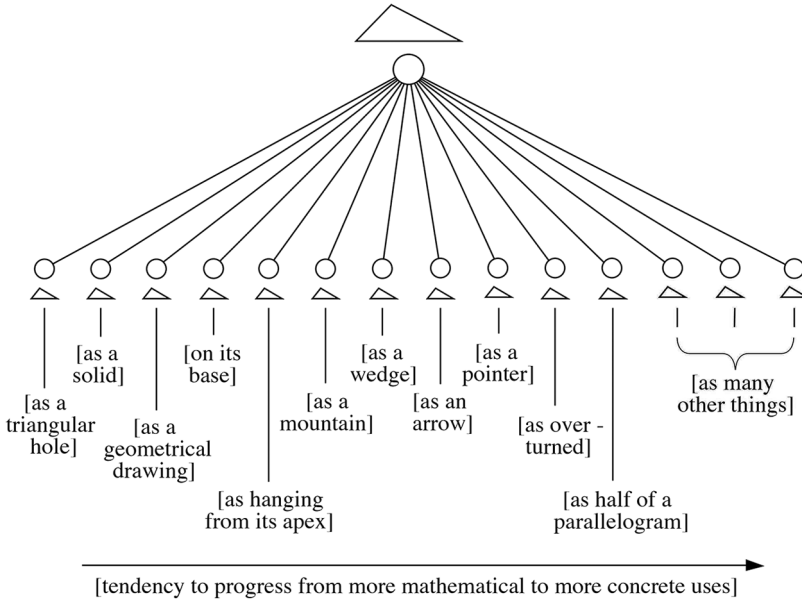
[as a
solid]   [on its
base]   [as a
wedge]   [as a
pointer]

[as a
triangular   [as a
hole]   geometrical   [as a   [as an   [as many
drawing]   mountain]   arrow]   other things]

[as over -
turned]

[as hanging
from its apex]   [as half of a
parallelogram]

[tendency to progress from more mathematical to more concrete uses]

FIGURE 2.   The diagram to visualize the multiple contexts (meanings) of Wittgenstein's
icon △ in Masterman, "Fans and Heads," p.45, reproduced by Tal Unreich.

patterns along the Wittgensteinian line of inquiry, Masterman transforms
the thesaurus into a philosophy machine (see "T," p. 195).

What her philosophy machine does is condense the whole of *Roget's The-
saurus* from one thousand semantic categories to eight hundred—which she
calls "heads"—and have them card punched laboriously for experiments on
Hollerith sorting machines.[50] This forms the basis of many of the computa-
tional innovations at CLRU, including the work of Spärck Jones. The latter's
doctoral thesis on *Synonymy and Semantic Classification* completed in 1964
pioneered the computational studies in information retrieval, and her influ-
ential paper on inverse document frequency established the basis for search
engines that would be adopted by Google.[51] These outcomes can be traced to
the initial decision at CLRU to use the thesaurus method, as Spärck Jones
explains:

> The CLRU addressed the problem of lexical disambiguation, and advo-
> cated the use of a thesaurus as a means of characterizing word mean-
> ings, in part because the structure of a thesaurus naturally supports
> procedures for determining the senses of words or, complementarily,

50. See Wilks, "Editor's Introduction," p. 7.
51. See Karen Spärck Jones, "A Statistical Interpretation of Term Specificity and Its Applica-
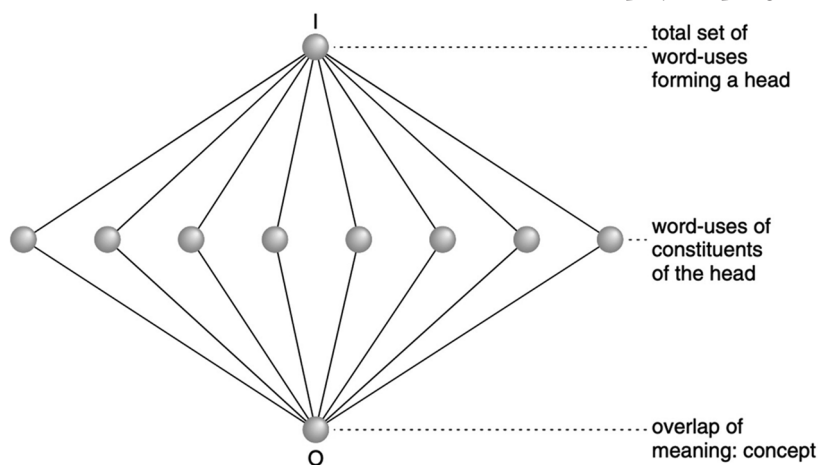tion in Retrieval," *Journal of Documentation* 28 (Mar. 1972): 11–21.

FIGURE 3.   The mathematical modeling of the head as an interpreted lattice, from Masterman, "Translation," p. 210.

for finding words for meanings. The assumption is that text has to be repetitive to be comprehensible so, in the simplest case in disambiguation, if a word's senses are characterized by several thesaurus classes, or heads, the relevant one will be selected because it is repeated in the list for some other text word. In text production, the fact that two heads share a word suggests that this is the right one.[52]

The distinction between "head" and "word" lies at the heart of the computable thesaurus. And what does that distinction signify? Spärck Jones refrains from putting her explanation in philosophical terms, whereas with Masterman—for whom everything must be worked out philosophically—the distinction of head and word is, first and foremost, a philosophical distinction between ideograph and logos. The computable thesaurus must rely on the head or the ideographic principle of the *zi* to embody many instances of family resemblances and semantic patterns in text production and speech contexts. The diagram above identifies the locations of head, word, and concept in a mathematical modeling of their relations (fig. 3). The head is the total set of word uses whereas the concept is the overlap of meanings (semantic aggregate) of the total set of word uses in the head. Masterman treats the head as ideographic sign, and equally important is the requirement that the basic units of calculation in the computable thesaurus be heads, not words.

This is because we are dealing with a thesaurus system that classifies words as a set of contexts rather than verbal synonyms. Each context is taken

52. Spärck Jones, commentary to Masterman, "Agricola in curvo terram dimovit aratro," in *Language, Cohesion and Form*, p. 159.

to be a single use of a word; as the sets of contexts can be infinite, there is indefinite extensibility of word meanings. The infinite sets of contexts are mapped onto a finite set of "heads" or eight hundred as adopted by CLRU. The context or word use in each of the heads falls into "lists" or "rows," where the "list" is a set of mutually exclusive contexts; the "row," on the other hand, is a set of quasi-synonymous contexts that can be used one after the other in an infinite string.[53] These elements are tagged in the numerical cross-reference system trackable by the computer. Take "839 LAMENTATION," where the numerical cross-references are interpreted as the overlap of meanings between the cross-referencing and cross-referenced:

839   *LAMENTATION*

Tag	*Word uses*
kind	lamentation, mourning;
	lament, wail, 363 INTERMENT;
	languishment, grief, moan, condolence, 915 CONDOLENCE;
	sobbing, crying, tears, mourning, 837 DEJECTION;
one be	sob, sigh, groan, moan;
	complaint, plaint, grumble, murmur, grief, 923 WRONG;
	mutter, whine, whimper, 886 CIVILITY;
bang kind	flood of tears, burst of tears, fit of tears;
	crying, howling, screaming, yelling, 411 CRY;
one bang be	spasm of sobbing, outburst of grief;
	cry, scream, howl, 411 CRY;
	wailing and gnashing of teeth, 900 RESENTMENT;
thing	weeds, crepe, crape, deep mourning, sackcloth and ashes, 225 INVESTMENT;
	passing-bell, knell, keen, death-song, dirge, 402 SOUND;
	requiem, wake, funeral, 998 RITE;
	[. . .]
	["T," p. 204]

53. See Masterman, "What Is a Thesaurus?" pp. 107–45.

The ellipsis refers to the remaining thirty-two rows of cross-referenced uses and repetitions that I cannot reproduce here. It is important to know that the ideographic head is indicated by capitalized letters—LAMENTATION, INTERMENT, CONDOLENCE, and so on—and each head is assigned a number to cross-reference with other heads like 839, 363, 915, 837, and so on. The ideograph is tagged and numbered in the series to classify a situation or action associated with other series of cross-referenced terms. The capitalized ideographic head LAMENTATION is distinguished from the word *lamentation* in the first row in the same manner as WARD is distinguished from the word *ward* in Masterman's extension of the *zi* principle to the English language.

To conclude, what does the machine modeling of the philosophical distinction between ideograph and logos accomplish? The answer is that the machine sets a contextually-based analysis of *semantic patterns* in motion; it means that the machine promises to measure up to the ambiguity and multiplicity of word meanings, analogies, metaphors, human creativity, and fallibility in language use; and it means that machine learning and AI technologies have been part of a major philosophical breakthrough even if their practitioners remain unaware of it. In that sense, the computable thesaurus is the philosophy machine Masterman and her team invented to bring the post-Wittgensteinian philosophy of language into being.